PB80-192495

Application of Photochemical Models. Volume I
The Use of Photochemical Models in Urban Ozone Studies

Association of Bay Area Governments
Berkeley, CA

Prepared for

Environmental Protection Agency
Research Triangle Park, NC

Dec 79

U.S. Department of Commerce
National Technical Information Service
NTIS

United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
Research Triangle Park NC 27711

EPA-450/4-79-025
December 1979

PB80-192495

Air

EPA

The Use of Photochemical Models in Urban Ozone Studies

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-450/4-79-025	2.	3. RECIPIENT'S ACCESSION NO. PB 80 192495
4. TITLE AND SUBTITLE Application of Photochemical Models in the Development of State Implementation Plans Volume I: The Use of Photochemical Models in Urban Oxidant Studies		5. REPORT DATE December 1979
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Wada, Ronald Y., et al.		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Association of Bay Area Governments Hotel Claremont Berkeley, California 94705		10. PROGRAM ELEMENT NO. 2AA635
		11. CONTRACT/GRANT NO. 68-02-3046
12. SPONSORING AGENCY NAME AND ADDRESS U. S. Environmental Protection Agency Office of Air Quality Planning and Standards Research Triangle Park, North Carolina 27711		13. TYPE OF REPORT AND PERIOD COVERED
		14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

This document describes procedures for application of photochemical models in the development of State Implementation Plans. Based largely on recent experience gained in photochemical model applications in the San Francisco Bay Area and in Denver, the guidance is directed toward potential model users in other ozone non-attainment areas. The guidance covers the following tasks: model selection; data collection and model input preparation including meteorological and topographical data, emission inventory data, ambient air quality data, treatment of initial and boundary conditions, and special field studies; the evaluation of photochemical model performance; model applications; and interpretation of model results with respect to attainment of the Federal ozone standard.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Photochemical Modeling SIP Development		

18. DISTRIBUTION STATEMENT RELEASE UNLIMITED	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE

EPA Form 2220-1 (9-73)

APPLICATION OF PHOTOCHEMICAL MODELS

Volume I

The Use of Photochemical Models in Urban Ozone Studies

prepared by

Association of Bay Area Governments
Hotel Claremont
Berkeley, California 94705

in association with

Bay Area Air Quality Management District
San Francisco, California

Lawrence Livermore Laboratory
Livermore, California

Systems Applications, Inc.
San Rafael, California

prepared for

U.S. Environmental Protection Agency
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina 27711

EPA Project Officer: John Summerhays

Contract No. 68-02-3046

Final Report, December 1979

PREFACE

This document is one of four volumes intended to provide information relevant to the application of photochemical models in the development of State Implementation Plans. The reports are particularly directed toward agencies and individuals responsible for preparation of non-attainment plans and SIP revisions for ozone. The four volumes are titled as follows:

Application of Photochemical Models

Volume I - The Use of Photochemical Models in Urban Ozone Studies

Volume II - Applicability of Selected Models for Addressing Ozone Control Strategy Issues

Volume III - Recent Sensitivity Tests and Other Applications of the LIRAQ Model

Volume IV - A Comparison of the SAI Airshed Model and the LIRAQ Model

This work is to a large extent based on the photochemical modeling experience gained in the San Francisco Bay Area in support of the 1979 Bay Area Air Quality Plan. The following individuals made significant contributions to this work:

Association of Bay Area Governments	- Ronald Y. Wada (Project Manager) - M. Jane Wong - Eugene Y. Leong
Bay Area Air Quality Management District	- Lewis H. Robinson - Rob E. DeMandel - Tom E. Perardi - Michael Y. Kim
Lawrence Livermore Laboratory	- William H. Duewer
Systems Applications, Inc.	- Steven D. Reynolds - Larry E. Reid

The authors wish to express their appreciation to John Summerhays, EPA Project Officer in the Source Receptor Analysis Branch of OAQPS, for his thoughtful review and comments on earlier drafts of this report.

TABLE OF CONTENTS

Page

1. INTRODUCTION 1
    BACKGROUND 1
    OVERVIEW OF THE MODELING PROCESS 3
        The Modeling Process for Ozone Plan Development . . . 4
        Schedule Considerations 8
        Resource Requirements 9
        Comparison of Rollback Modeling Concepts with Deterministic Modeling Concepts 12
    REFERENCES 16
2. MODEL SELECTION 19
    SELECTION CRITERIA 19
    THE SELECTION PROCESS 21
    GUIDELINES FOR SELECTION OF GENERAL MODELING APPROACH . . . 22
        Urban Areas with Extensive Monitoring Networks 24
        Urban Areas with an Average Number of Monitoring Stations 25
        Urban Areas with Limited Monitoring Networks 25
3. INPUT DATA COLLECTION AND MODEL INPUT PREPARATION 27
    METEOROLOGICAL AND TOPOGRAPHICAL DATA 30
        The Prototype Day Approach 30
        Assessing Data Requirements 31
        Sources of Data 32
        Field Studies 33
        Data Preparation and Preprocessing 33
        LIRAQ AND SAI Model Applications 34

TABLE OF CONTENTS (Continued)

Page

EMISSION INVENTORY DATA 35

Grid Selection 36

Overview of the Inventory Effort 37

Spatial and Temporal Distribution of Stationary Source Emissions 39

Temporal Distribution of Stationary Source Emissions 41

Stationary Source Projection Methods 42

Motor Vehicle Emissions 44

Hydrocarbon Species Allocation 45

AMBIENT AIR QUALITY DATA 47

TREATMENT OF INITIAL AND BOUNDARY CONDITIONS 52

Methods for Treating Initial Conditions 55

Base Year 55

Future Year Simulations 56

Boundary Conditions 57

Examples of the Treatment of Boundary Conditions in Baseline Calculations and Model Evaluation Studies 58

Future Year Simulations 60

Summary 61

SPECIAL FIELD MEASUREMENT STUDIES 62

Meteorological Data 63

Air Quality Data 63

Collecting Supplementary Field Data for LIRAQ 64

CALTRANS Field Studies in Los Angeles 66

REFERENCES 67

TABLE OF CONTENTS (Continued)

Page

4. THE EVALUATION OF PHOTOCHEMICAL MODEL PERFORMANCE 69
    - OUTLINING THE MODEL EVALUATION STUDY 71
        - Definition of the Size and Boundaries of the Modeling Region 71
        - Definition of Spatial and Temporal Model Resolution 71
        - Time Period to be Simulated 72
        - Choice of Meteorological Conditions 72
        - Determination of the Number of Meteorological Regimes 73
        - Performance Measures and Standards 74
    - MODEL EVALUATION PHASE 76
        - Adaptation of the Model to the Study Area 76
        - Collection of Model Inputs 77
        - Performance of Model Simulations 77
        - Assessment of Model Simulation Results 77
        - Rectification of Inadequate Performance 79
    - EXAMPLES OF PREVIOUS MODEL EVALUATION STUDIES 82
        - The LIRAQ Model 82
        - The SAI Airshed Model 83
    - REFERENCES 90

5. MODEL APPLICATIONS 93
    - BASELINE PROJECTION 93
    - EMISSIONS SENSITIVITY ANALYSES 94
    - CONTROL STRATEGY SIMULATIONS 94
    - ALTERNATIVE PROGRAM DESIGNS 94

Page

PRACTICAL ASPECTS OF DESIGNING A MODEL APPLICATION STUDY 96

MODEL APPLICATIONS IN THE SAN FRANCISCO BAY AREA 99

Baseline Forecasts 99

Emissions Sensitivity Analyses 99

Control Strategy Simulations 99

MODEL APPLICATIONS IN THE DENVER METROPOLITAN AREA 105

Baseline Forecasts 105

Sensitivity Analyses 105

SUMMARY OBSERVATIONS 110

REFERENCES 111

6. MODEL INTERPRETATION 113

THE "WORST CASE" ISSUE 113

IMPERFECT MODEL PERFORMANCE 114

SPECIFICATION OF INITIAL AND BOUNDARY CONDITIONS 116

METHODS FOR MODEL INTERPRETATION USED IN THE SAN FRANCISCO BAY AREA 116

METHODS FOR MODEL INTERPRETATION USED IN THE DENVER METROPOLITAN AREA 119

CONCLUDING REMARKS 119

APPENDICES

APPENDIX A PREPARATION OF METEOROLOGICAL INPUT FIELDS FOR LIRAQ SIMULATIONS IN THE SAN FRANCISCO BAY AREA

APPENDIX B PREPARATION OF METEOROLOGICAL INPUT FIELDS FOR SAI SIMULATIONS IN DENVER AND LOS ANGELES

APPENDIX C PREPARATION OF EMISSION INVENTORIES FOR PHOTOCHEMICAL MODELING IN THE SAN FRANCISCO AND DENVER REGIONS

LIST OF FIGURES

PAGE

Figure 1-1 The Air Quality Implementation Planning Process 5

Figure 1-2 Sample schedule for modeling and related plan development tasks for preparation of a 1982 SIP revision 10

Figure 3-1 Map of Bay Area showing the region within which 5-km grid resolution is available in LIRAQ. 38

Figure 3-2 Process for preparation of separate, disaggregated trip end and hot stabilized motor vehicle emission inventories. 46

Figure 3-3 Schematic diagram of grid cell inputs. 53

Figure 4-1 Examples from the comparison of available station observations of oxidant with LIRAQ calculated time histories of ozone 84

Figure 4-2 Sample scatter diagrams of observed versus calculated hourly average NO2 concentrations, observed versus calculated station mean concentrations, and observed verus calculated station maximum concentrations from the LIRAQ simulations. 85

Figure 4-3 Examples of observed and SAI Airshed model predicted hourly ozone concentrations (pphm) at various stations in Denver on 28 July 1976 88

Figure 4-4 Estimate of bias in SAI Airshed model predictions as a function of ozone concentration. 89

Figure 5-1 Sample diagram of potential emissions sensitivity test cases 95

Figure 5-2 Sample schedule for model application tasks. 97

Figure 5-3 Example LIRAQ results--2000 baseline ozone projections. 101

Figure 5-4 Plots of unadjusted and adjusted regionwide and high hour ozone as a function of % reductions of 1985 HC emissions. 102

List of Figures Continued

PAGE

Figure 5-5 Control strategy testing with the AQMP modeling system 103

Figure 5-6 Reduction in predicted ozone concentrations (pphm) at Denver stations due to predicted emissions changes. 108

Figure 6-1 Sample statistical extrapolation to determine "worst case" ozone in a future year 115

Figure 6-2 LIRAQ verification for two 1973 prototype days using 1975 emissions, based on 16 hourly values at 15 locations. 118

LIST OF TABLES

PAGE

Table 1-1 Estimated resource requirements for photochemical model application. 13

Table 2-1 Partial inventory of photochemical models. 23

Table 3-1 Example levels of detail in data used as input to photochemical models 28

Table 3-2 Typical variables for emission projection for the Bay Area 43

Table 3-3 Summary of LIRAQ compatible percent emissions by source type. 48

Table 4-1 Sample photochemical model performance measures and standards. 75

Table 4-2 Summary of SAI experience in evaluating the performance of the SAI Airshed model 86

Table 5-1 Baseline LIRAQ projections for the San Francisco Bay Area 100

Table 5-2 LIRAQ emission sensitivity analysis results. 100

Table 5-3 Summary of control strategies tested in the San Francisco Bay Area 106

Table 5-4 Effectiveness of alternative control strategies for the San Francisco Bay Area 107

1. INTRODUCTION

BACKGROUND

The purpose of this document is to provide technical information on the use of advanced photochemical models (primarily Eulerian grid models and secondarily Lagrangian Trajectory models) in the development of State Implementation Plans. Both Lagrangian trajectory models and Eulerian grid models are mathematical representations of the physical and chemical processes that occur in the atmosphere to form ozone*. As their respective names imply, Lagrangian trajectory models simulate ozone production within a parcel of air as it follows a prescribed trajectory, while Eulerian grid models simulate ozone production and transport in all cells of a regular square or rectangular grid.

Ozone levels well in excess of the National Ambient Air Quality Standard have been a persistent and pervasive problem. Many of the control measures potentially needed to attain the ozone standard across the nation are complex, expensive, and controversial. In contrast, previously applied techniques for determining control requirements for precursor emissions are generally acknowledged as unvalidated and of limited applicability. Thus, there is a need for making available to State and local air quality planning agencies improved analytical techniques for evaluating the effectiveness of alternative ozone control measures.

The history of photochemical modeling has been split into two distinct areas: research and model development, and model applications. The state-of-the-art in ozone model development has been summarized by many individuals in varying contexts. The 1976 International Conference on Photochemical Oxidant Modeling and Its Control contained several papers in this area, most notably that by Demerjian (1976). Roth, et al. (1976) have also performed an in-depth evaluation for the American Petroleum Institute. Ozone modeling has achieved a high level of complexity and sophistication. While there continue to be potential problem areas and uncertainties regarding the suitability of existing ozone models for addressing specific issues such as long range transport, with proper application and interpretation of results, their use in SIP analyses is now encouraged by EPA.

---

*In this document reference is made to both ozone and oxidants. Ozone is the primary component of photochemical oxidants in most areas. Advanced photochemical models simulate ozone production, and the current EPA reference method for oxidant monitoring measures ozone. Previously used monitoring methods and models (e.g., EPA's "Appendix J" rollback model) were based on oxidants. The recent change from a 0.08 ppm (1-hr) oxidant standard to a 0.12 ppm (1-hr) ozone standard places the standard, the monitoring method, and advanced models on the same basis.

In stark contrast, the application of models in support of ozone plan development has until recently been restricted to the most primitive forms available. Legislative mandates and the need to develop ozone control plans have created a demand for improved analytical tools to assess the effects of alternative control strategies on ozone levels.

Summerhays (1976) has inventoried ozone model applications as of that date. Those previous applications have not been conducted to systematically assess alternative ozone control strategies. Since that time, the Systems Applications, Inc. (SAI) Urban Airshed Model has been applied in the Denver metropolitan area to support the evaluation of oxidant control strategies, and the Livermore Regional Air Quality (LIRAQ) model (see MacCracken and Sauter, 1975) has been applied in direct support of the development of an oxidant control plan for the San Francisco Bay Area. As summarized by Wada, Leong, and Robinson (1977), LIRAQ was used to guide the selection of control strategies to attain and maintain the (former) federal oxidant standard in the region. A number of problem areas were encountered during the course of the analysis and were subsequently overcome. Guiding the modeling effort was an interagency modeling committee including modeling experts from a broad spectrum of participating agencies. The combined expertise of the modeling committee made it the forum for discussion and resolution of problems related to model applications. Successful completion of the analysis was in large part attributable to the functioning of the committee.

Much of the guidance presented here takes advantage of the modeling resources, experience and process developed for the Bay Area Air Quality Maintenance Plan.

The Environmental Protection Agency has been involved over the past few years in the development of guidelines for applying air quality models in varying contexts:

- Evaluating indirect sources (U.S. EPA, 1975a)
- Applying atmospheric simulation models to air quality maintenance areas (U.S. EPA, 1975b)
- Background information on the procedures, uses, and limitations of oxidant prediction relationships (U.S. EPA, 1977 and 1978a)
- Guideline on Air Quality Models (non-photochemical) (U.S. EPA, 1978b)
- Procedures for comparing air quality models (U.S. EPA, 1978c)

This document details the use of photochemical models in the development of ozone control strategies for State Implementation Plans and includes procedural descriptions directed toward agencies responsible for preparation of such plans. Particular emphasis is placed in the document on:

- planning the total effort
- collecting and preparing meteorological, topographical, air quality, and emissions data
- assigning initial and boundary conditions
- evaluating model performance
- practical aspects of model application
- interpreting model results

In each area, general discussion is followed by specific examples of previous model applications to illustrate how a particular problem has actually been treated. Time and resource requirements are also noted throughout the document to assist other agencies in developing the modeling program most appropriate to the problems and constraints in their regions.

The information is addressed to individuals in planning and/or control agencies who have a basic familiarity with commonly existing emission inventory and meteorological data bases, rudimentary modeling techniques (e.g., EPA's UNAMAP models), and commonly accessible computer facilities. The document covers those tasks which should be common to most regional photochemical modeling efforts. Specific machine-related procedures or other problems that are unique to each modeling effort must be addressed individually by each model user.

OVERVIEW OF THE MODELING PROCESS

Models are representations of our current understanding and information regarding a process. For photochemical oxidants (ozone), models are representations of current knowledge and data regarding the origins of precursor emissions, their transport and chemical transformation, and the resultant distribution of pollutants in space and time. Modeling facilitates the evaluation of complex problems through the analysis of each sub-component of the problem. These components are then assembled to study the problem in its entirety. Models are thus vehicles for quantifying complex relationships, such as are embodied in the formation of ozone and other oxidants.

Air quality models always have and will continue to contain inherent inaccuracies. This is a reflection of practical limitations of data acquisition, computer capacity, and an incomplete understanding of important atmospheric processes. However, air quality problems exist now; to make rational control decisions which directly address ambient air quality standards established to protect public health and welfare, no viable alternatives can be identified to the use of air quality models. The more relevant and meaningful questions are: How will modeling results be translated into emission control requirements, and does that procedure provide flexibility and opportunities for "mid-course correction?"

The air quality implementation planning process is schematically summarized in Figure 1-1. The need for air quality models is a direct consequence of the definition of air quality goals in terms of ambient standards. Models serve as the bridge between ambient standards and source control requirements; they may be broadly defined as expressions of the relationship between source emissions and ambient air quality.

A second feature illustrated in Figure 1-1 is the continuing feedback and reevaluation of control strategies required as part of the planning process. This feature is required in recognition of the dynamic nature of both air pollution problems and our understanding of how to deal with them. The key point is that it is not necessary to specify the ultimate solution to the air quality "problem" via a single analysis undertaken at a single point in time. Indeed, the process has been designed with the recognition that it is not realistic to do so. This concept is fundamental to the air quality planning process.

The Modeling Process for Ozone Plan Development

The process for applying a deterministic model in support of a plan may be divided into several discrete steps, as follows:

- Model selection - This step consists of two parts; first, deciding whether to use a photochemical dispersion model; and second, selecting a specific model. All of the information presented in this document is in part intended to assist in the decision of whether to use a photochemical model by describing what would be involved in such an effort. The decision should be based on the availability of adequate time and resources and on the need for credibility that a photochemical dispersion model generally provides. Section 2 contains specific model selection criteria as well as a step-by-step selection procedure.

- Data collection and model input preparation - This step consists of:

    - assessing the existing data base and identifying inadequacies in meteorological and topographical data, emission inventory data, and ambient air quality data

    - planning and conducting a field study to collect supplementary air quality and meteorological data

    - planning and developing suitable emission inventories (spatially and temporally disaggregated, speciated, and projected to future years)

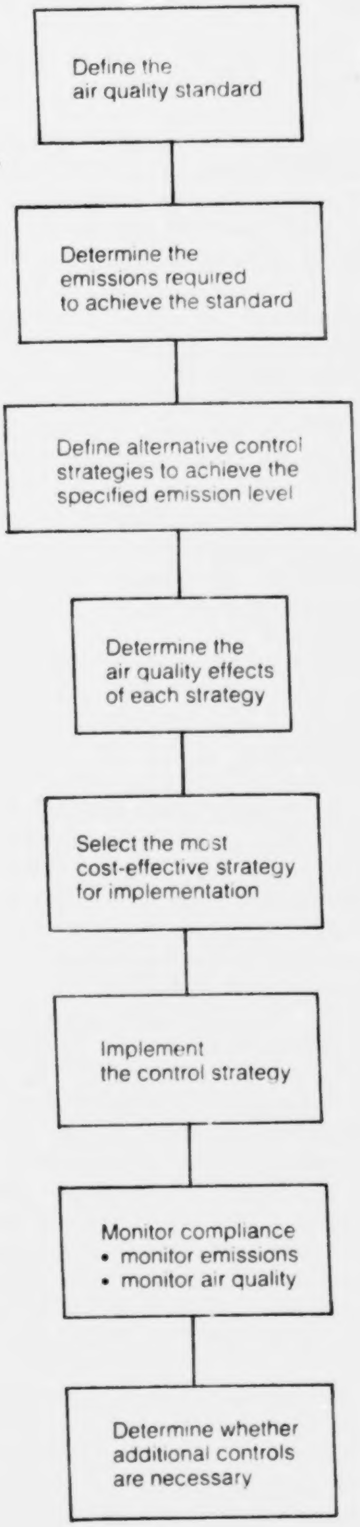

Define the air quality standard
Determine the emissions required to achieve the standard
Define alternative control strategies to achieve the specified emission level
Determine the air quality effects of each strategy
Select the most cost-effective strategy for implementation
Implement the control strategy
Monitor compliance
• monitor emissions
• monitor air quality
Determine whether additional controls are necessary

Figure 1-1

The air quality implementation planning process

- preparing the input data in model-compatible format

Section 3 and Appendices A, B, and C contain detailed information regarding each of these steps.

- o Model performance evaluation - This step consists of adapting the model to the study area, performing initial model simulations based on the input data collected, assessing the simulation results, and rectifying inadequacies in model performance. These steps are reviewed in detail in Section 4.

- o Model applications - There are a variety of model applications that may be made depending on the specific issues to be addressed. Section 5 describes the types of applications that are useful in the development of an air quality plan. In addition, a variety of ozone control issues of potential interest are described in Volume II of this report.

- o Model interpretation - Three significant issues regarding the use and interpretation of modeling results to demonstrate the adequacy of a given control strategy for meeting the ozone standard are discussed in Section 6.

Careful planning of the total modeling effort will help the user to make the best use of available resources. The user must identify the scope of the effort, the model to be used, desired accuracy and acceptable model limitations, the existing emissions and aerometeric data for preparing inputs, and the resources available for collecting supplemental data. First a general plan should be formulated to define the anticipated applications, how model performance will be measured, and what constitutes adequate model performance. Although these expectations may change during the course of the project it is important to establish pertinent goals at the outset in order to provide some basis for making the planning decisions, especially those relating to data collection and model modification.

After a general plan has been developed, the data requirements are examined in light of the available data base in order to ascertain the need to collect additional aerometric or emissions data. The planning stage must occur far enough in advance of the actual model simulations to allow the planning and execution of any special monitoring program that may be necessary to collect supplemental field data. In general, the requisite amount of input data depends on the model chosen, the region and conditions of interest, and the desired model performance. If NEDS emission inventory data are being considered for use, that data should be carefully checked and updated as necessary to ensure that the data conforms to the input requirements of the model to be used.

Once the input files have been constructed and the model has been appropriately modified, the first simulation is carried out. The results from this simulation are analyzed as described in Section 4. If model performance is judged unacceptable, then either the input data must be modified (perhaps by reanalyzing the available data or collecting additional data) or the model must be modified and the simulation rerun. This process is repeated until satisfactory results are obtained. Additional evaluation runs will not only supply information on how the model performs under different conditions but can also point out weaknesses in the data or the model. If suitable performance cannot be achieved, it may be necessary to select another model or to make an assessment of the consequences of using an improperly verified model in the intended applications study.

Once adequate model performance has been verified, model applications to address the control issues related to the air quality plan may proceed. All previous steps have been conducted in preparation for this phase. Here, it is important to have clearly identified the issues to be addressed, and how the model is to be employed in each case. Modifications to input file (i.e., emission inventory files, initial and boundary conditions) to simulate the effect of various controls should be thoroughly documented, and duplicate files stored as a contingency against the need to re-create the simulations at a future time (e.g., during the review of the plan).

Of crucial importance in the execution of each step is the necessity for interagency participation and peer review of the assumptions, procedures, and judgments made. There are two factors that contribute to this necessity. First, as previously mentioned, air quality models always have and will continue to contain inherent inaccuracies, such that perfect performance cannot be expected. This means that the way in which model results are interpreted can substantially influence subsequent control decisions. Second, models tend to be highly complex, such that they are viewed suspiciously by non-modelers as being "black boxes" over which there is very little control. Under these conditions, it is desirable to establish the credibility of the results by setting up an open process by which technical staff from a variety of organizations can participate in making the variety of technical assumptions, judgements, and interpretations which must invariably be made along the way.

The participation and review process should be accompanied by the preparation of formal, detailed documentation at each step. This contributes to the openness of the process, facilitates review by interested parties, and can minimize any subsequent controversy regarding the technical basis of the resulting plan.

In the San Francisco Bay Area, the interagency participation and review process is centered in a "modeling committee". The committee meets roughly once per month (or more often, as necessary) to review and discuss various aspects of the modeling work as they progress, and also functions as a vehicle for coordinating the transfer of data between agencies. Modeling experts from the following organizations participate on the committee:

- o Association of Bay Area Governments (lead non-attainment planning agency)

- Bay Area Air Quality Management District (primary user of the LIRAQ photochemical model)
- Metropolitan Transportation Commission (responsible for transportation inputs to the modeling analysis)
- Lawrence Livermore Laboratory (authors of the LIRAQ model)
- California Department of Transportation
- California Air Resources Board (State agency that must review and adopt the plan)
- Environmental Protection Agency, Region IX (EPA regional office that must review and accept the plan)
- Systems Applications, Inc. (a private photochemical modeling firm)

All technical memoranda and issue papers related to the modeling work are reviewed in draft form by this committee before being widely disseminated.

It should be understood from the beginning that this type of open process substantially reduces efficiency in the preparation of the modeling analysis. (However, such a process should also minimize any technical disagreements that might otherwise occur during the plan review period.) Technical objectivity and a spirit of cooperation among the participating agencies are required so that analyses can be completed in a timely manner.

Schedule Considerations

Of the five steps in the modeling process previously listed, the collection and preparation of input data will generally be the most time consuming. No clear relationship between the quality and quantity of input data, and the quality of modeling results has yet been established. The quality or credibility of a given result is largely a subjective assessment at this time*. Therefore, considerable latitude exists in deciding how much time and monetary resources to devote to the modeling analysis as a whole, and to each step of the analysis. If the ozone problem in a given area is of such a magnitude that the control strategies to be considered are expected to be costly and controversial, then it is crucial that the modeling analysis supporting the plan be as complete and rigorous as possible.

---

*EPA is sponsoring efforts to develop systematic model performance evaluation methods, as well as additional analyses to determine the sensitivity of model performance to input data of varying quality and quantity. See Section 4 and Volume III of this report for additional discussion.

Figure 1-2 is a sample schedule for preparation of a modeling analysis for a 1982 State Implementation Plan revision. This schedule is presented to illustrate the phasing of tasks in relation to one another. Actual time required for any specific application may vary from that shown depending on a number of factors, i.e., the time required, if any, to collect supplementary field data, the number of cases to be simulated and the expected turnaround time at the computer facility, etc. It is important to establish the boundaries of the schedule in the context of the development of the total plan. Estimates should be made for time required to prepare the plan, state and local plan review, modification and approval, and submission to EPA. Depending on the significance of the controls recommended and the institutional complexity of the region, the plan review and approval process may range from 6 months to well over a year.

Resource Requirements

Successful application of a photochemical air quality simulation model requires personnel with several skills and fields of expertise. The discussion below is not meant to imply that a large number of people are required; indeed, in some instances, only a few are all that are needed if they possess the requisite experience and understanding.

Project Manager. Some form of overall project management or supervision is required for coordination of activities such as data acquisition and verification, codification and inspection of model inputs, computer simulation, analysis of results, and reporting. The project manager should have an understanding of all facets of the air quality modeling program, including an appreciation of how the results of the study are likely to be interpreted and used.

Meteorologist. Able assistance in the field of boundary layer meteorology is an indispensable part of the modeling program. The main responsibility of the meteorologist is to determine which of the available algorithms for computing wind fields, mixing depths, and diffusivity fields are most appropriate for the application at hand. In some cases, it may be advantageous for the meteorologist to develop new algorithms for preparation of the meteorological inputs. In addition, he must check the resultant meteorological inputs for accuracy and consistency with physical intuition.

Atmospheric Chemist. Knowledge of atmospheric chemistry is central to the modeling effort. An atmospheric chemist should be called upon as necessary to estimate hydrocarbon species present as initial and boundary conditions. Also, the chemist is involved in the analysis and interpretation of the simulation results.

Computer Specialist. The computer specialist must be thoroughly experienced with the computer system on which the model is to be implemented--typically one of the large-scale computer systems such as the IBM 360/370, CDC 6600-7600, Univac 1108/1110, or Honeywell 6000 series. For any particular system, the specialist must be familiar with the system design, data file storage, tape handling procedures,

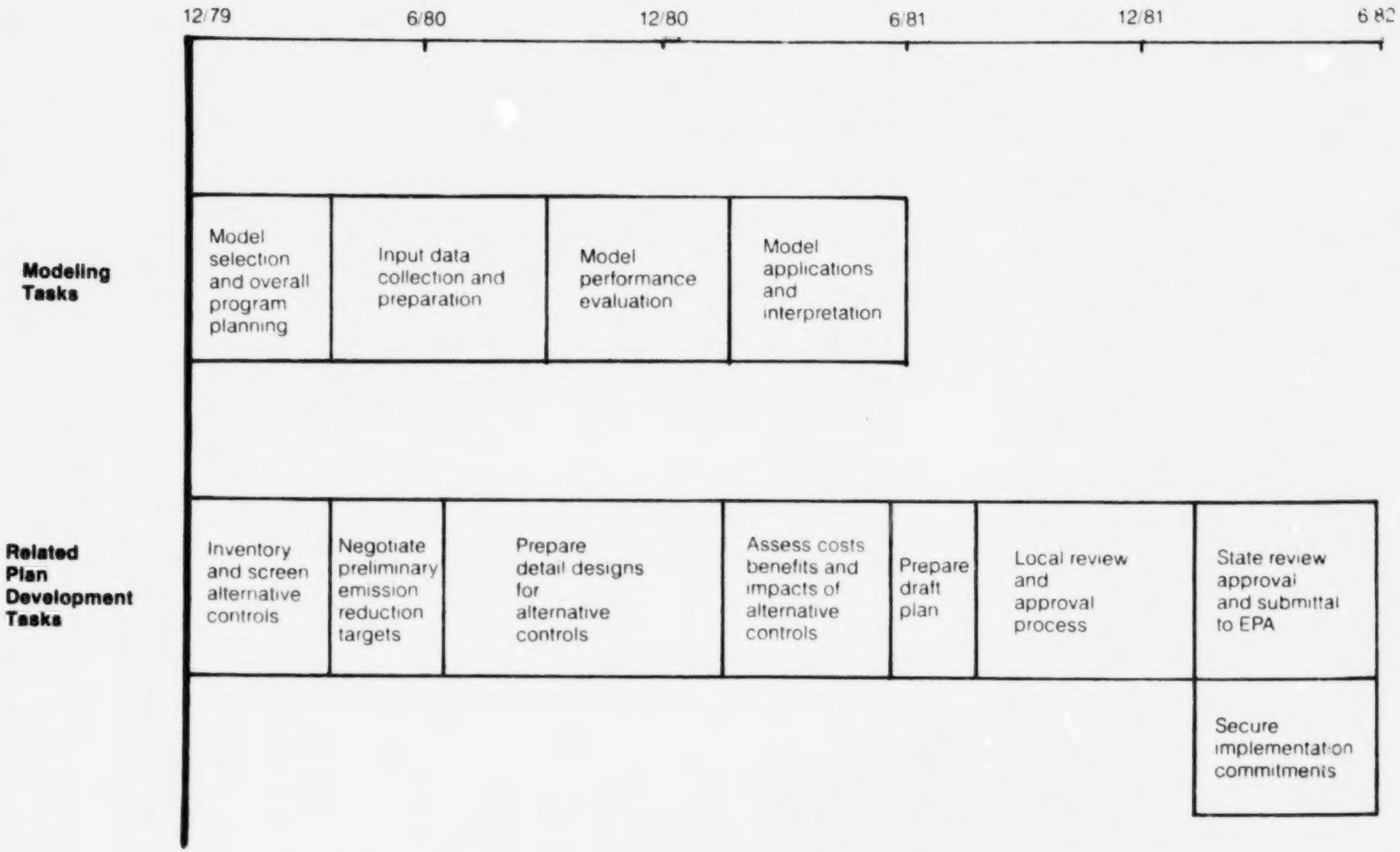


Figure 1-2

Sample schedule for modeling and related plan development tasks for preparation of a 1982 SIP revision

structured programming, and the appropriate program language. Additionally, the specialist should be familiar with computer plotting hardware and software (if such is available) because much of the model output is more easily understood when presented graphically.

Air Pollution Control Engineer/Emission Inventory Specialist. Familiarity with the source inventory being used, its weaknesses and inherent assumptions in its derivation are also of importance. This expertise is of particular importance when simulating the effects of alternative control strategies.

Data Analyst. The data analyst has two main responsibilities. First, he inspects the accuracy and consistency of the raw air quality, meteorological, and emissions data to be used in the data preparation programs. This person also participates in the review of the model inputs prior to their use in the simulation. Thus, it is highly desirable that the analyst have an understanding of meteorology, atmospheric chemistry, and air quality modeling. Second, he interprets the modeling results. Although some of the models do provide software for generating graphical displays and output statistics, it is possible that additional analysis of the simulation results will be desired.

Technical Illustrator. Because the results of the modeling analysis are likely to be distributed formally, it may be worthwhile to prepare graphs or diagrams presenting certain of the more significant results of the study. In that case, the services of a technical illustrator may be desirable.

Labor and computing costs for the modeling effort are estimated on the basis of past experience acquired by the organizations participating in this study. The estimates necessarily presume that the individuals performing the modeling analysis already possess a sound understanding of atmospheric photochemical modeling in general and of the specific intended model in particular. In the absence of this experience, additional time will most likely be required to acquaint the user with the model and to develop the requisite level of experience in its use. The ranges in the estimates are rather large for several reasons, including:

- Variability in the skill and experience of the individuals using the model.
- Variability in the complexity of the application (e.g., unusual meteorological or emissions patterns).
- Variability in the scope and depth of the proposed emission control strategies.
- Variability in the amount of existing data (e.g., transportation patterns, emissions, growth projections, air quality, and meteorology) and the extent to which existing data have been consolidated into a modeling data base.
- Variability in the time available with which to plan and carry out the modeling study.
- The extent to which the study area has been modeled previously.

- The required accuracy of the modeling analysis.

Table 1-1 summarizes the estimates of the ranges in level of effort (in weeks) required to perform the various tasks in the modeling program. In estimating these figures, we assume that model performance would be evaluated for two different meteorological scenarios. Also included are estimated ranges in the computing time required to accomplish the various tasks assuming the usage of a grid model, such as LIRAQ or the SAI Airshed Model.*

The variability in the ultimate cost of preparing a model for use depends largely on whether an adequate data base exists for the urban area. When adequate data are available, model preparation and performance evaluation can normally be carried out for about $75,000 to $150,000. However, if an emissions inventory must be compiled and if a short term intensive field program is mounted to collect additional air quality and meterological data, then the overall costs may run as high as $500,000 to $750,000. We note that the cost of even more extensive supplemental emissions inventory and aeromatric monitoring programs (such as EPA's Regional Air Pollution Study in St. Louis) can exceed several million dollars. In view of the resources required to perform the modeling analysis, thorough planning of all phases of the modeling program is indispensable.

Comparison of Rollback Modeling with Deterministic Modeling

There are several key differences between deterministic photochemical models (i.e., Lagrangian trajectory models and Eulerian grid models) and rollback models which, if not understood at the outset, could lead to confusion later in the planning process.

Rollback models require as input aggregated regional emission inventories for a baseyear and future years, and a "design value" highest expected oxidant measurement in the base year. With this information, the percentage control of hydrocarbon emissions required to attain the standard is computed based on the ad hoc presumption of a linear relationship between hydrocarbon emission reductions and reductions in maximum ambient ozone levels.

The technical basis for rollback models suffers from major deficiencies. Among those deficiencies are the following:

- The role of NOx emissions in the formation of oxidants is ignored.
- Varying photochemical reactivities for different organic compounds are ignored.

---

*Computer times assume usage of CDC 7600 computer.

TABLE 1-1. Estimated Resource Requirements for Photochemical Model Application

Activity	Level of Effort (weeks)	Estimated Cost for Labor (dollars)	Computing Requirements* (hours)
PROGRAM DESIGN			
Episode Selection	2-4	2,400-4,800	--
Control Strategy Review	1	1,200	--
Identification of data requirements	1	1,200	--
Prescription of model evaluation procedures	1-2	1,200-2,400	--
DATA ACQUISITION			
Routine aerometric data	1-2	1,200-2,400	--
Readily available emission inventory	2-4	2,400-4,800	--
Development of a new inventory	50-200	60,000-240,000	1-2
Special studies--aerometric and emissions inventory	1-2	1,200-2,400	--
Acquisition of new aerometric data	100-250	250,000-500,000	--
Data base management	6-24	7,200-28,000	1-3
INPUT DATA PREPARATION			
Mixing depths	4-6	4,800-7,200	--
Wind fields	4-24	4,800-28,000	--
Solar radiation	2	2,400	--
Initial and boundary conditions	8-12	9,600-14,400	--
Hydrocarbon speciation	1	1,200	--
Emissions	4-6	4,800-7,200	--
Miscellaneous	4-6	4,800-7,200	--
Total computing resources needed for data preparation	--	--	2-6
MODEL EVALUATION	3-16	3,600-19,200	3-8
MODEL APPLICATION	5-40	6,000-48,000	2-10
MODEL INTERPRETATION	6-16	7,200-19,200	0.5
REPORTING	8-12	9,600-14,400	--
TOTALS**	64-625	84,000-948,400	10-30

---

*Assuming the usage of a CDC 7600 Computer.

**The lower figures are for areas with extensive existing data bases that are easily converted to appropriate model inputs; the higher figures are for areas where extensive additional data must be collected to provide a good data base for model application.

- The air quality effect of control strategies which result in non-uniform emission reductions over a given region cannot be evaluated.

- Rollback models are difficult to validate or verify.

The modeling process associated with deterministic models is considerably more complex than the rollback procedure. Such models compute ozone concentrations for selected days on an hour-by-hour basis at specific locations. They have been designed to directly address those fundamental characteristics of the ozone problem which are ignored by rollback models:

- The role of NOx is explicitly described in a chemical kinetic model.

- Differing reaction rates for different classes of hydrocarbon compounds are explicitly described.

- Spatial and temporal aspects of the problem and of alternative controls can be evaluated.

- Partial verification is possible by simulating a number of days in the historical record.

Key concepts associated with the use of deterministic models are as follows:

The Prototype Day Concept - Deterministic models are mathematical representations of the physical processes which lead to the formation and transport of pollutants with time scales ranging from a few minutes to a few days. This level of resolution means that instead of the base year required by rollback, a base or prototype day is of greater significance. The performance of a deterministic model is determined by simulating several historical days in order to verify that it can reproduce the ozone levels that were measured on those days. Having been verified on specific days, it is then necessary to assume that the models can forecast pollutant concentrations under different emission scenarios using the same meteorological conditions which occurred on the prototype days.

The basic assumption is that the meteorological conditions which lead to adverse levels of air pollutants are recurring phenomena which are not expected to be changed in the future by long term trends. While this assumption is true in the sense that high pollution days have common characteristics, it may also be expected that no future day will ever be precisely the same as a past day. Since the precise characteristics of future days cannot be predicted, it must be assumed that the historical day(s) used are reasonably representative of what will occur in the future.

The Worst Case Concept - The worst case concept is the result of interpretations of ambient air quality standards, and the rollback models were designed to be responsive to this requirement. In attempting to apply the concept to the more complex modeling process for deterministic models a number of difficulties arise.

First, there must be a mechanism for relating the historical day or days selected for modeling to worst case conditions in the the future. From a technical viewpoint it is extremely difficult to predict the conditions which will lead to worst case ozone levels, especially if future emission patterns will be significantly changed from existing emission patterns. The quantity and distribution of pollutant emissions in a given area may change dramatically over time as a result of implementation of control programs, and overall growth patterns. Those conditions which led to worst case ozone levels in the past may not lead to worst case levels in the future.

Second, the differences between "worst" case, "second-worst" case, and "third worst" case may not be solely due to differences in meteorological conditions. For example, these differences may be due to quirks in the emission inventory for specific days. These factors are difficult to document and account for, hence the inventory used is usually a typical weekday emission pattern.

Third, the worst case for one monitoring location is not necessarily the worst case for other locations. To consider worst case conditions at all locations would require many more validation days (at least one for each location), with a concomitant spiraling of time and budget requirements.

These difficulties are not encountered under the rollback procedure because the rollback model is not sensitive to the various factors of concern. In Section 6, methods for addressing this problem are discussed.

The Emissions Target Concept - Because the output of deterministic models is in terms of air quality parameters (concentrations of pollutants) there is no need to define an intermediate "emissions target." The effects of alternative control strategies on air quality can be tested directly, and because the effects of NOx and differing hydrocarbon reactivities are included, it is probable that a number of different emission levels will result in the same maximum ozone levels. Thus, the concept of a single emissions target developed under the rollback procedure is no longer applicable. Emissions targets can be expressed in a variety of ways (different mixes of precursor emissions and different spatial and temporal patterns), each of which could result in attainment of the ozone standard.

It should be clear from the preceeding discussion that a subtle change in the concepts and perspectives regarding the role of models in the planning process must occur if a deterministic model is to be applied. Those previous concepts that are unique to rollback modeling have tended to become imbedded into the way many individuals perceive the modeling process for ozone, and should be discarded if a deterministic model is to be used.

REFERENCES

Association of Bay Area Governments, Bay Area Air Pollution Control District, and Metropolitan Transportation Commission, "Draft Air Quality Maintenance Plan," Appendix C - Results of the LIRAQ Emissions Sensitivity Analysis, December 1977.

Demerjian, K. L., "Photochemical Air Quality Simulation Modeling: Current Status and Future Prospects," International Conference on Photochemical Oxidant Pollution and Its Control, September 12-17, 1976, Raleigh, North Carolina, Proceedings: Volume II, January 1977, EPA-600/3-77-016.

MacCracken M.C., and G.D. Sauter, "Development of an Air Pollution Model for the San Francisco Bay Area, Final Report to the National Science Foundation," Volume I, October 1, 1975, Lawrence Livermore Laboratory, UCRL-51920.

Roth, P.M., et al., "An Evaluation of Methodologies for Assessing the Impact of Oxidant Control Strategies," prepared for American Petroleum Institute, August 24, 1976, Systems Applications, Inc., EF76-112R.

Summerhays, J. E., "A Survey of Applications of Photochemical Models," International Conference on Photochemical Oxidant Pollution and its Control, September 12-17, 1976, Raleigh, North Carolina, Proceedings: Volume II, January 1977, EPA-600/3-7-016.

Wada, R. Y., E. Y. Leong, and L. H. Robinson, "A Methodology for Analyzing Alternative Oxidant Control Strategies," *Journal of the Air Pollution Control Association*, Vol. 29, No. 4, pp. 346-351, April 1979.

U.S. Environmental Protection Agency, "Evaluating Indirect Sources," Volume 9 of *Guidelines for Air Quality Maintenance Planning and Analysis*, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, Revised, September 1978, EPA-450/4-78-001. (a).

U.S. Environmental Protection Agency, "Applying Atmospheric Simulation Models to Air Quality Maintenance Areas," Volume 12 of *Guidelines for Air Quality Maintenance Planning and Analysis*, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, 1975. (b)

U.S. Environmental Protection Agency, "Uses, Limitations and Technical Basis of Procedures for Quantifying Relationships Between Photochemical Oxidants and Precursors," Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, EPA-450/2-77-021a, November 1977.

U.S. Environmental Protection Agency, "Procedures for Quantifying Relationships Between Photochemical Oxidants and Precursors: Supporting Documentation," Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, EPA-450/2-77-021b, February 1978. (a).

U.S. Environmental Protection Agency, "Guideline on Air Quality Models," Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, EPA-450/2-78-027, April 1978. (b).

U.S. Environmental Protection Agency, "Workbook for Comparison of Air Quality Models," Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, EPA-450/2-78-028a, and EPA-450/2-78-028b (appendices), May 1978. (c)

2. MODEL SELECTION

The first step in the process for applying a photochemical model is model selection. This step should ordinarily precede the collection and preparation of input data, despite the fact that most deterministic photochemical models have similar data requirements. Differences in input data requirements and format from one model to another could result in added costs and schedule delays if the input data is prepared first.

SELECTION CRITERIA

In considering the selection of a model, the potential user must weigh the advantages of greater credibility and capability against the disadvantage of greater cost, time and personnel requirements. Specific considerations important when deciding on a model are:

Desired Capabilities

- Physical and chemical phenomena of importance - The model should ideally be capable of simulating the physical and chemical phenomena that are known or suspected to be important determinants of the ozone problem in the region.
- Previous model evaluation results - Most of the models that are available have been previously applied to one or more urban areas. Review of past performance may provide valuable insight with regard to the weaknesses of a given model, and will also provide some perspective on what may be expected for model performance.
- Control strategies to be tested - If it is known in advance what control strategies are to be tested, this may have a bearing on the model selected. For example, if selective control of hydrocarbon species is contemplated, or if relative effectiveness of elevated point source versus ground level area source emissions control is an important issue then certain models are better equipped than others to address the issues of concern. Volume II contains additional information in this regard.

Resource Constraints

- Availability of resources to collect appropriate input data - Input data requirements will vary from one model type to another and from one region to another. For example, Eulerian grid models require a more complete specification of the flow field than Lagrangian trajectory models, and produce a more complete representation of air quality across a region as a result. A given region may exhibit a greater or lesser degree of variability in its flow field depending on topography,

Preceding page blank

presence of bodies of water, and other factors. The greater the complexity the more data required to represent the conditions. Thus, input data availability and the potential for additional data acquisition are important considerations in model selection.

- Cost - The primary costs associated with the application of a photochemical model are due to input data collection and preparation, model execution, and diagnosis of model problems. In each case, both staff and computer resources can be significant. Resource limitations can effectively preclude application of the more sophisticated models.

- Computer constraints - The accessibility of appropriate computing facilities can be an important consideration. For example, it would be difficult to use the LIRAQ model on any computer other than the CDC 7600. The current version of the SAI model is operating on CDC and UNIVAC systems. Some program modifications would be necessary to implement the model on other computer systems. Other models may have comparable constraints. In addition, many government-operated computer systems are not set up to accommodate the substantial computing demands of photochemical models. As a minimum, the conversion of any advanced model to operate on a computer system other than the system on which it is currently operated will involve additional costs and time delays.

- Personnel constraints - The capabilities of the model user in terms of future model applications may be important. Since an in-house capability to use the model in future applications is generally desirable, the potential for existing or future personnel to be trained in model application, troubleshooting and model interpretation may be an important consideration in deciding on the degree of complexity of the model to be used.

- Schedule constraints - Schedule constraints are usually dictated by regulations specifying deadlines for submittal of plans. A tight schedule may preclude the conduct of field studies, may place substantial pressure on the preparation of input data, and may force added expenditures for computer time and consultants.

The criteria listed are not independent of one another. For example, deficiencies in input data could be overcome by conducting special field studies. Such studies tend to be expensive, time consuming, and may involve costs comparable to or greater than the model application effort. Consultants can be used to supplement existing user staff, or in lieu of user staff in order to minimize personnel and schedule constraints; again at some additional cost. Selection of an appropriate computer facility can make a substantial difference in "turnaround time" for model runs, but can also mean added costs.

Conversely, it is often the case that resources for a given plan development effort are fixed in advance, and thus the feasibility of implementing a given model may be limited by the magnitude of resources allocated.

THE SELECTION PROCESS

As previously stated, model selection involves weighing the user's needs against available resources. In general, grid models provide the greatest capability, credibility, and are the most expensive to apply; EKMA is least expensive and least credible; and trajectory models are in between. (See Volume II of this report for additional discussion.) The process for selection of a model consists of a number of specific steps:

(1) Identify the physical and chemical phenomena that are either known or suspected to be important determinants of the ozone problem in the region (e.g., typical wind flow patterns during high ozone episodes, significant sources of background ozone, presence and behavior of elevated inversion layer(s), presence of photochemical aerosol, geographic extent of high ozone levels, etc.).

(2) Identify available air quality, meteorological, and emissions inventory data--number and location of monitoring stations, length of historical record, and ease of access and processing of such data by computer.

(3) Identify (if possible) important aspects of potential control strategies that might require specific modeling capabilities. (See Volume II for discussion of examples such as selective hydrocarbon species control, long range transport, elevated vs. low-level emissions control, etc.).

(4) Inventory alternative computer facilities available, including the potential for use of large remote facilities (it is typical of governments to require a great deal of red tape and prior approval before any non-government operated computer facility can be used).

(5) Evaluate resource constraints for the total model application effort.

(6) Estimate the minimum number of model runs needed, including model performance evaluation and strategy evaluation runs (detailed suggestions for this estimate are given in Section 5).

(7) Prepare a preliminary schedule for plan development to obtain a sense of time available for each task. This schedule may consist of guesses, but will still be a valuable exercise when evaluating the feasibility of application of alternative models.

(8) Identify the most promising models through literature review. The models identified are to be subjected to more intense investigation.

(9) Contact EPA and/or the developers of the models selected for further investigation to determine:

- o availability of the code and user's manual
- o recent updates/modifications
- o computer requirements
- o other questions that may have arisen during the literature review

If it is anticipated that extensive use of consultants will be made, this contact may take the form of a request for proposals and subsequent proposal evaluation.

(10) Array the advantages and disadvantages of each model with respect to the selection criteria. It is likely that no model will be able to satisfy all of the criteria and some tradeoff decisions will be necessary (e.g., a typical tradeoff is cost versus technical precision).

(11) Select the model.

Table 2-1 is a partial inventory of alternative photochemical models, indicating the model developer, where it has been previously applied, and appropriate literature references. Additional information is contained in Volume II. The contact point within EPA for information regarding the models is:

Chief, Source Receptor Analysis Branch
Monitoring and Data Analysis Division
Office of Air Quality Planning and Standards
Environmental Protection Agency
Research Triangle Park, North Carolina 27711
(919) 541-5391

GUIDELINES FOR SELECTION OF GENERAL MODELING APPROACH

For urban areas, the appropriateness of a given modeling approach is often most constrained by data availability. For the purpose of this discussion, three levels of extent of monitoring networks providing data are defined: extensive, average, and limited. Examples of such categories of monitoring networks are:

Table 2.1. Partial Inventory of Photochemical Models

MODEL	DEVELOPER	PREVIOUS APPLICATIONS	REFERENCES
EKMA	U.S. EPA, OAQPS Research Triangle Park, N.J.	Many applications, no documented verification available	o U.S. Environmental Protection Agency, "Uses, Limitations, and Technical Basis of Procedures for Quantifying Relationships Between Photochemical Oxidants and Precursors," Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina, November 1977, EPA-450/2-77-021a. o U.S. Environmental Protection Agency, "User's Manual for Kinetics Model and Ozone Isopleth Plotting Package," Office of Air Quality Planning and Standards, Research Triangle Park, North Caroline, July 1978, EPA-600/8-78-014a.
Langrangian Trajectory			
ELSTAR	Environmental Research & Technology, Inc. Boston, Massachusetts	St. Louis	o Lloyd, A.C., et. al., "The Adaptation of a Lagrangian Photochemical Air Quality Simulation Model to the St. Louis Region and the Regional Air Pollution Study Data Base," Environmental Research and Technology, Inc., prepared for U.S. Environmental Protection Agency, Meteorological Assessment Division, Research Triangle Park, North Carolina, December 1978.
TRACE	Pacific Environmental Services, Santa Monica, CA	Los Angeles	o Drivas, P.J. and L.G. Wayne, "Validation of an Improved Photochemical Air Quality Simulation Model," Pacific Environmental Services, Inc., Pub. No. TP-016, Santa Monica, California, March 1977. o Drivas, P.J., "TRACE (Trajectory Atmospheric Chemistry and Emissions) User's Guide," Pacific Environmental Services, Inc., Pub. No. TP-016, Santa Monica, California, November 1977.
SAI TRAJECTORY	Systems Applications, Inc. San Rafael, California	Sacramento	o Meyers, T.C., et al., "User's Guide to the SAI Trajectory Model," Systems Applications, Inc., San Rafael, California, SAI Report No. EF 79-3, prepared for U.S. Federal Highway Administration, January 8, 1979.
Eulerian Grid			
LIRAQ	Lawrence Livermore Laboratory, Livermore, CA	San Francisco St. Louis	o MacCracken, M.C., et al., "The Livermore Regional Air Quality Model: I. Concept and Development," Journal of Applied Meteorology, Vol. 17, pp. 254-272, March 1978. o Duewer, W.H., et al., "The Livermore Regional Air Quality Model: II. Verification and Sample Application in the San Francisco Bay Area," Journal of Applied Meterology, Vol. 17, pp. 273-311, March 1978.
SAI AIRSHED	System Applications, Inc. San Rafael, California	Los Angeles Denver St. Louis Sacramento	o Reynolds, S.D., L.E. Reid, and T.W. Tesche, "An Introduction to the SAI AIRSHED Model and Its Usage," Systems Applications, Inc., San Rafael, California, December 1978, SAI Report No. EF 78-53R.
IMPACT	Science Applications, Inc. La Jolla, California	Sacramento	o Fabrick, A., et al., "Point Source Model and Development Study," Science Applications, Inc., Westlake Village, California, prepared for California Air Resources Board and the California Energy Resources Conservation and Development Commission, March 1977. (Appendix C is the User's Guide to IMPACT.)
MADCAP	Science Applications, Inc. La Jolla, California	San Diego	o Sklarew, R.C., "Verification of the MADCAP Model of Photochemical Air Pollution in the San Diego Air Basin," presented at the American Meteorological Society Annual Meeting, Reno, Nevada, January 1979.

Monitoring Network	Urban Area
Maximum	St. Louis
Common	Houston, Philadelphia
Minimum	Denver

These different levels of data base quality are described in more detail in Section 3.

In the discussion that follows, guidelines are suggested for the selection of an appropriate modeling approach for use in different types of situations. For each situation, the general spectrum of measurements that might comprise the data base are described as well as the factors of importance that might influence the model selection process--such as the likelihood of importing high concentrations of precursors, relative contributions of natural hydrocarbon sources to pollutant loading, and potential for characterizing flow patterns in the area. Recommendations are then made concerning the modeling approach that might be most suitable for use in the particular application.

Urban Areas with Extensive Monitoring Networks

For areas where substantial resources for data collection are available (or where extensive supplementary data collection has already occurred), it is appropriate to select one of the Eulerian grid models that includes photochemistry for an evaluation of a control strategy. The use of such a model offers several advantages. It permits one to examine the effects of changes in:

- o The spatial and temporal distribution of NOx emissions, hydrocarbon emissions, or both;
- o The reactivity of the hydrocarbon emissions;
- o The background concentrations;
- o The contaminants transported into the area.

If the control strategy focuses on emissions from an isolated or a small area source, it may be advantageous to apply a trajectory-based (Lagrangian) model; this type of model is comparatively easy to apply, and the results it generates are likely to be less expensive and as appropriate as those obtained using a grid-based model. However, trajectory models are best applied in areas free from significant horizontal variations in the concentration field, and in which there are no significant wind shear effects. If the proposed control strategy would be likely to cause uniform changes in the regional emissions distribution, then EPA's EKMA model should be considered as well. In general, however, Eulerian grid models should be considered for use as the primary evaluation tool.

Urban Areas with an Average Number of Monitoring Stations

A less extentisve and more common data base compilation effort will result in a lower level of confidence in the accuracy of the modeling results. Nevertheless, with the use of judicious estimates for unavailable input information (or ranges in the input variables), meaningful concentration estimates can be made using either grid or trajectory models. In situations where only the grid or trajectory models contain the detail required for the strategy evaluation, these models are a particularly important means of estimating the impacts on precursor and ozone concentrations of transport, complex topography, or proposed changes in the spatial distributions of emissions.

As the availability of data decreases and the physical character of the areas and control strategies become simpler, use of the EPA EKMA isopleth model becomes more attractive. If the strategy considers control of only an isolated point source or of emissions in a limited area, a reactive plume model or a trajectory model would be a possible candidate for use. If the impact of transport and changes in the spatial distribution of emissions were relatively unimportant, the EKMA model would again be a candidate. EKMA would also be appropriate if the proposed emissions changes were spatially and temporally uniform, with no change in hydrocarbon reactivity.

Urban Areas with Limited Monitoring Networks

Urban areas with limited monitoring networks and for which supplementary data are not collected are likely to experience significant problems in attempting to evaluate control strategies. To assess the effects of either the changes in the spatial distribution of emissions or the reactivity of hydrocarbon emissions in regions that exhibit complex flow behavior or variable upwind concentrations, a grid or trajectory model should be used. The data required for the evaluation and exercise of these models are greater in both breadth and quantity than the data normally available. In a case where the control strategy to be evaluated includes neither severe changes in the spatial distribution of emissions nor changes in the reactivity of hydrocarbon emissions, and where the area has relatively uncomplicated flow patterns and upwind boundary conditions, then the EKMA model would be appropriate.

3. INPUT DATA COLLECTION AND PREPARATION

Deterministic models require a substantial amount of input data in order to simulate the production and transport of ozone and its precursors. The purpose of this section is to outline the various types of data required and to describe specific considerations in the preparation of the data for model use. The data have been divided into three categories: meteorological and topographical data; emissions inventory data; and air quality data. Meteorological and topographical data are used to describe the flow and dispersion of pollutants; emissions inventory data are the principal inputs that are varied to test the effectiveness of control alternatives in reducing ambient pollutant concentrations; air quality data are used to determine initial and boundary conditions for the model and to evaluate model performance. This section concludes with guidance in two areas related to input data preparation: the specification of initial and boundary conditions, and the conduct of special field studies.

The information is presented on a general level, since each model will have unique requirements for input data format that are more appropriately contained in their user's guides. Emphasis has therefore been placed on acquainting the potential user with the types and quantities of data required, potential sources of data, and the significance of the data to the modeling effort. Table 3-1 delineates three example levels of detail in data for photochemical modeling. These levels represent data that are prepared for input to a photochemical model, and may include data from both existing monitoring systems and supplemental field data. The "maximum practical level" corresponds to the most extensive data base currently available or potentially available given present funding constraints and the state of the art in photochemical modeling. In many respects, the RAPS (Regional Air Pollution Study) data base in St. Louis is an example of this category. A data base with a "minimum acceptable level" of detail might be adequate for some modeling purposes, but numerous assumptions would have to be invoked in preparing model inputs. For example, to estimate mixing depths over a city where upper air temperature soundings are unavailable, one might assume that the vertical temperature gradient measured at some nearby city reflects conditions in the city of interest. Between these two levels of detail lies the "commonly-used level," which includes those data bases that have been employed in previous photochemical modeling studies. This does not suggest, however, that such a data base is well suited to model evaluation and application. Some of the measurements generally lacking or in short supply in such data bases may be ones to which model performance is quite sensitive. At both the minimum and commonly used levels, a field observation program to supplement the available data base is highly desirable for improving model performance and minimizing modeling uncertainties. Specific judgments regarding the nature of additional data to be collected will depend heavily on the resources available and the existing data base in each region; Table 3-1 may be useful as an aid in identifying data needs in each specific case.

Preceding page blank

TABLE 3-1. EXAMPLE LEVELS OF DETAIL IN DATA USED AS INPUT TO PHOTOCHEMICAL MODELS

Input	Maximum Practical Level	Commonly Used Level	Minimum Acceptable Level*
Mixing depth	Continuous monitoring of mixing depths with acoustic sounder at one or more locations Several (5-20) vertical temperature soundings throughout the day at various locations within the modeling region Numerous surface temperature measurements recorded hourly at various locations throughout the modeling region One or more instrumented towers providing continuous measurements of the mixed layer thermal structure	A few (2-5) temperature soundings at different times of the day at one or two locations Several surface temperature measurements recorded at various locations throughout the modeling region	Single daily temperature sounding at an airport within or nearby the region being modeled A few (1-3) surface temperature measurements with which to estimate temporal variation
Wind fields	Numerous ground-based monitoring stations reporting hourly average values Frequent upper air soundings at several locations throughout the modeling region Continuous upper level measurements on one or a few elevated towers Wind, inversion, temperature, and terrain data used as input to a 3-D numerical model yielding a mass conserving wind field	Interpolation from ground-based monitoring network and limited (1-5) number of upper level soundings at one or two locations Resultant wind field rendered mass consistent by divergence-free algorithm	Interpolation from limited (2-5 stations) routine surface wind data; theoretically derived vertical profile assumed
Solar radiation	Several (5-10) UV pyranometers located in the region, continuously recording UV radiation levels Vertical attenuation of radiation at a few locations several times daily determined by aircraft observations Spatial (3-D) insolation fields determined by interpolation of measurements	One or two UV pyranometers; insolation assumed constant over the region Vertical attenuation estimated empirically as a function of aerosol mass	No radiation measurements available; estimated theoretical values based on the solar zenith angle Attenuation not accounted for
Boundary and initial conditions	Hourly species concentrations extrapolated and interpolated throughout the region using data from the extensive ground-based monitoring network; airborne data also available; hydrocarbon mix obtained from gas chromatographic analyses at several times during the day	Hourly concentrations extrapolated and interpolated using data from several ground-based stations; hydrocarbon mix obtained from gas chromatographic analysis at one or two stations one or a few times during the day or on similar days, limited airborne data available	Hourly concentrations extrapolated and interpolated from a minimal routine monitoring network; either hydrocarbon mix assumed or average value obtained from a compilation of available data taken in a similar area No data on concentration variations aloft

Input	Maximum Practical Level	Commonly Used Level	Minimum Acceptable Level*
Stationary source emissions	Separate gridded inventories for point and area stationary sources; characterization of organic composition, and NO/NO2 emissions rates for major sources; diurnal and seasonal variations in nominal emissions rates for each major source type based on questionnaires or stack sampling	Lumped, gridded inventory for stationary sources; no species fractionation; seasonal and diurnal variation in regional emissions for each pollutant	Lumped stationary source emissions inventory for the region as a whole; limited information on the percentage of each source type; no temporal variation
Hydrocarbon species distribution of emissions	Mix obtained from gas chromatographic analysis of samples collected throughout the region, particularly near large sources or from stack samples	Mix obtained from standard emissions factors (AP-42) together with a detailed source inventory, supplemented with one or two gas chromatographic analyses	Mix assumed or obtained from available data compilation, either for the city of interest or some similar area
Mobile source emissions	MOBILE-1 emissions factors used in conjunction with local vehicle age distribution; corridor-by-corridor VMT, including peak and off-peak speed distributions, vehicle mix, and traffic data for intrazonal trips Spatial and temporal distributions of cold starts inferred from actual traffic and demographic data Cold start factors applied grid by grid when calculated mobile source emissions	MOBILE-1 emissions factors, assumed vehicle mix, and intrazonal VMT; estimated peak and off-peak speeds, fewer traffic counts available for verification, VMT available for fewer major arterials Cold starts temporally resolved using traffic distribution; no spatial resolution or spatial resolution only from estimates of driving patterns	Gridded VMT, emissions factors estimated from 49 state mix, and average (FDC) driving profile; assumed regional speed distribution Cold starts as a fixed percentage of all driving--traffic data are not detailed enough for spatial resolution of cold starts; cold starts estimated from demographic data
Air quality data for evaluating model performance	Hourly averaged species concentrations for NO, NO2, O3, NMHC, CO, and particulates from an extensive ground-based monitoring network	Hourly averaged concentrations of NO, NO2, O3, NMHC, CO, and particulates from several ground-based stations	Hourly averaged concentrations of NOx, O3, THC, and CO from a minimal routine monitoring network

---

*Using data at this level of detail necessitates numerous assumptions.

METEOROLOGICAL AND TOPOGRAPHICAL DATA

The meteorological data requirements of air quality models vary widely depending on the sophistication of the model and the application. For example, for some applications of a simple Gaussian plume model, only a single surface wind velocity and an estimate of stability may be required. In more complicated applications, information on the depth of the mixed layer and wind values updated at specified time intervals may be required. Lagrangian models, which follow a parcel of air, generally require more detailed information on surface flow fields. Finally, the Eulerian (grid-based) models require the most sophisticated time- and space-dependent meteorological input fields.

This document is primarily concerned with the grid-based models. The meteorological data collection and preparation for these models involves:

- assessing model requirements for accuracy and spatial and temporal resolution of meteorological variables,
- locating data sources,
- planning and conductng supplemental field studies where required,
- acquiring the topographic, wind, temperature and other meteorological measurements,
- selecting prototype meteorological conditions (normally from the set of days for which supplemental field data are available),
- selecting the gridded region,
- generating properly formatted inputs for preprocessors, and
- examining output fields and modifying the inputs to the preprocessors in an iterative manner until the output fields are thought to be realistic.

The Prototype Day Approach

For air quality planning we attempt to simulate days in a recent base year on which high ozone levels were observed. Once satisfactory model performance is achieved for such days we then assume that the model can forecast ozone concentrations for different future emissions scenarios under the same meteorological conditions that occurred on the validation days. This assumption is based on the expectation that a) similar adverse meteorological conditions will re-occur in the future, and b) those conditions that now produce high ozone levels will do so in the future. Ideally, one would like to identify the adverse conditions associated with future emissions patterns; however, this is difficult at best to accomplish based on current methods.

Preparation of meteorological inputs for sophisticated grid-based photochemical models is very labor-intensive, time-consuming and therefore expensive. For example, preparation of meteorological input files for simulation of one 24-hour period for LIRAQ in the San Francisco Bay Area may require up to three months of effort of an experienced meteorologist. Therefore, simulations are ordinarily performed on a small number of prototype periods ranging in duration from several hours to a few days. Because of the expense, it is important to select very carefully those prototype days that are to be used in modeling.

The selection process often involves making some trade-offs between the desirability of a day for simulation and the amount of meteorological information available for that day. For example, if field studies were performed in a base year, they may not have included the day with the highest or second highest, or most widespread ozone levels during that year. (Because the ambient air quality standard for ozone is interpreted in terms of the worst days, it would be desirable to simulate the worst or second worst day of a given year.) The meteorological data base for the worst or second worst day may not be sufficiently detailed to ensure adequate model performance. In such cases a less desirable day with a good data base would be preferred, provided that there is subsequent analysis to determine a worst case scenario as discussed later in Section 6.

Assessing Data Requirements

One of the first steps in planning air quality simulations is to assess meteorological data-gathering and processing requirements. This step is crucial: first, because decisions will have a large impact on the cost of the project (e.g., how large and expensive a field study is necessary?), and second, if data requirements are misjudged, the model may fail to produce acceptably realistic simulations.

The grid-based models require time-dependent, two or three-dimensional fields of wind velocity and mixing height throughout the simulation period. These fields are specified by inputing limited surface wind and temperature observations and even more sparse wind and temperatures aloft into special preprocessing programs. Because the space/time density of wind and temperature measurements is always limited, various assumptions must be made about the structure of the atmosphere aloft for input to the preprocessors. Preprocessors are designed to produce physically consistent time-dependent mixing height fields and wind velocities for each grid cell in the modeled region.

The required density, in time and space, of meteorological inputs to the preprocessor programs depends upon: a) the computational requirements of the codes needed to perform adequate interpolation, extrapolation or smoothing from limited available measurements, b) the geographical complexity of the region to be modeled, c) the complexity of atmospheric circulations on those days chosen for simulation, and d) the model's sensitivity to variations in the meteorological variables. In a region with flat, relatively uniform terrain a comparatively sparse network of observations may be adequate to characterize the region's wind and mixing height fields. On the other hand, in a geographically complex region like the San Francisco Bay Area, where winds are channeled by terrain and are subject to land/sea breeze reversals, and where land/sea effects produce large horizontal gradients in mixing height, a large body of observational information must be collected, analyzed and formatted so that preprocessors can produce realistic characterizations of the wind and mixing layer dynamics.

Initial assessments of meteorological data requirements should be made by a group of individuals with expertise in a) the functioning and input requirements of the model and preprocessor codes, b) local air quality, c) the small- and mesoscale local atmospheric circulations -- particularly during air pollution episodes, and d) any past or ongoing local research or field studies that might yield insights into the nature of wind and stability fields. Such a group would be competent to address such questions as:

- o What meteorological conditions characterize local air pollution episodes?
- o Which days should be chosen for simulations?
- o What grid network is needed to track the pollutant cloud?
- o What is the reliability of local meteorological measurements? (including evaluation of instrument exposure and averaging times)
- o Is the network of local surface observing stations sufficiently dense?
- o Are data from the nearest upper-air measurement station representative of the area?
- o Is the available data base, together with knowledge of local conditions, adequate to prepare model inputs or are additional field measurements required?

Sources of Data

Topographical data can be obtained from the United States Geological Survey (U.S.G.S.). Average elevations of specified grid compartments can be determined from U.S.G.S. contour maps. Data for the San Francisco Bay Area for LIRAQ were taken from a U.S.G.S. California topography tape that gave mean elevations for 1 x 1 minute sub-quadrants that were interpolated to a 1 km grid.

Surface measurements of meteorological variables may be available from a variety of sources - at least in the metropolitan centers. Local meteorologists should be consulted when performing a survey of available local data. Surface measurements of wind velocity and temperature are often collected at major airports and military installations, by air pollution control agencies, colleges and universities, by various Federal, state and local governmental agencies and by some public utility and other private companies. Weather data from many sites throughout the United States are placed in archives by and available from:

National Climatic Center
Environmental Data and Information Service
National Oceanic & Atmospheric Administration
Federal Building
Asheville, NC 28801

The scarcity (in time and space) of upper-air wind and stability measurements is usually the most troublesome problem facing the analyst who is attempting to prepare meteorological input fields to sophisticated grid-based photochemical models. The National Weather Service (NWS) operates a nationwide network of upper-air stations that are spaced approximately 250 to 500 miles apart. Soundings of wind, temperature and humidity are made twice daily, at 12:00 and 00:00 GMT. In many areas the NWS upper-air network is the only source of scheduled observations aloft. In some areas, however, routine upper-air measurements (from pilot balloons, radiosondes, instrumented aircraft, towers, or acoustic sounders) may be made by universities, some governmental regulatory agencies or research organizations.

Field Studies

When the available routine meteorological observations in a given region are judged to be inadequate to construct two- or three-dimensional, time dependent wind and mixing-height fields, special field observation programs may be needed. Typically such observations, designed to fill the most important gaps in the existing observation network, include supplemental measurements of winds and temperatures aloft obtained from pibals, radiosondes, instrumented aircraft and acoustic sounders. They should also include additional surface wind, temperature, and solar radiation measurements in the more data-sparse or topographically complex areas. Occasionally a field experiment is performed to study a specific phenomenon; an example of this was the fluorescent particle tracer experiment conducted in the Bay Area to study vertical mass transport across the inversion interface (MacCracken and Sauter, 1975). Further discussion of field studies is presented later in this section.

Data Preparation and Preprocessing

After prototype days have been selected and all the meteorological and air quality observations have been gathered, several preprocessing tasks must be performed to convert the observations into a format compatible with model input requirements. First, it will probably be desirable to key some of the data into computer data files to facilitate error-checking, conversion of units, preparation of inputs for preprocessor codes and for input to post-processor (model performance evaluation) codes. A detailed discussion of the data management system that was set up for LIRAQ applications in the Bay Area can be found in MacCracken, 1975.

Grid-based models generally require that wind velocity, mixing layer depth and other variables (e.g., solar flux density) be specified for each grid cell in the modeling region. Obviously it is not practical to measure the variables in each cell, so one or more preprocessors are employed to compute values for each cell from relatively sparse observations. Preprocessors generally perform interpolation, extrapolation and, sometimes, smoothing. They may also calculate wind and other quantities aloft based on a few observations and specified assumptions regarding the vertical distribution of wind and stability. The choice of a particular preprocessing scheme will depend upon the structure and complexity that one wants to attribute to the meteorological fields. More specific comments on preprocessors used for the LIRAQ and SAI models are given in the next section.

Once a preprocessing methodology is decided upon, the necessary input fields must be prepared from the set of observations. Interpolation schemes often tend to give unrealistic values in data-sparse portions of the grid. In order to prevent this, it is sometimes necessary to supply fictitious values at some locations. The required number, positioning and magnitudes of the fictitious values are largely determined by experience. It may be necessary to modify the inputs and rerun a preprocessor several times before satisfactory output fields are obtained -- fields that are ready for input to the air quality model.

LIRAQ and SAI Model Applications

The methodologies for preparing meteorological input fields for grid-based models will vary because models are different and because each region that is to be modeled is unique. Thus, although general guidance can be given, specifics must be defined in the course of the actual model application in a particular area. In order to provide insights into some of the problems that are encountered and how they are treated, the following discussion highlights some of the methods and assumptions used in LIRAQ simulations in the San Francisco Bay Area and in the SAI Urban Airshed Model applications in Denver and Los Angeles.

The selection of prototype days in all three regions relied heavily on those days for which supplemental meteorological field data were available. Prototype day selection often involves deciding whether to simulate a day with "ideal" conditions (such as the day with the year's highest or second highest ozone concentration) or a day with more complete meteorological information.

The choice of the gridded modeling region also involves tradeoffs. Because of computer limitations, the number of grid cells that can be handled is limited. LIRAQ simulations in the Bay Area were performed with a 20 x 20 grid of 5 km elements. In some of the SAI simulations in Denver and Los Angeles, each region was subdivided into 900 two-mile squares. The modeling region must always be selected carefully to ensure that important source and receptor areas are included, that boundary effects are minimized, and that flow reversals on a particular prototype day do not transport the pollutant "cloud" out of the region and then return it later.

Preprocessing of inversion base height fields is best tailored to the region that is to be modeled. Overall, the preprocessing of these fields was relatively uncomplicated in the SAI/Denver studies and complex in the LIRAQ/Bay Area studies. For example, in the Denver simulations the depth of the mixed layer was assumed to be uniform in space. In the Los Angeles studies the depth of the mixed layer was assumed to be a function of distance from the ocean. In the Bay Area, because of the influences of the ocean, bays and adjacent mountain ranges, relatively complicated inversion base height fields were constructed.

Methodologies for generating wind fields can also be tailored to a specific region. In the Denver, Los Angeles, and Bay Area simulations surface wind observations were fed into an interpolation scheme to obtain a field of surface wind values throughout the gridded region. In Denver, winds aloft were set equal to surface values in one set of simulations. In later simulations a more sophisticated approach that accounts for local convergence or divergence effects was used. In the Los Angeles studies three different methods were used to calculate the three-dimensional wind field. Two of these produced reasonable results. In the Bay Area the MASCON mass-consistent atmospheric flux model was employed (Dickerson, 1978). The MASCON code was also used to generate two-dimensional fields of atmospheric transmissivity from Eppley pyranometer measurements taken at nine locations. Appendix A presents a more detailed summary of the meteorological inputs used in the LIRAQ/Bay Area simulations. Appendix B presents the same kind of example information for the SAI studies in Denver and Los Angeles. These detailed examples are included to enhance the reader's understanding of how the data have been prepared in the past, and are not necessarily intended as guidance for future efforts.

EMISSION INVENTORY DATA

A multitude of guidelines documents, computerized systems, and reference works have been prepared to assist the development of emission inventories. The most recent and relevant guidance is being published by EPA under the title:

Procedures for the Preparation of Emission Inventories
for Volatile Organic Compounds
Volume II Emission Inventory Requirements for
Photochemical Air Quality Simulation Models
(EPA-450/4-79-018)

That report contains the basic guidance necessary for preparation of emission inventories for photochemical models. The material presented in this section is intended to complement previously published guidance in this area by focusing on the aspects of inventory preparation that are of key importance to planning applications of photochemical models.

The general problem for source inventory builders for photochemical models is to develop procedures and techniques to adapt an existing source inventory to photochemical model format. As a minimum, this requires a change from the normal basis of tons/day, annual average, by county, to a basis of gm/sec, hourly average, by grid square for a specific day. In general, it should not be assumed that an adequate inventory already exists.

The purpose of this section is to summarize general procedures and techniques for spatial and temporal distributions and projection methods. Emphasis is focussed on the San Francisco Bay Area experience, and the following areas are covered:

- Grid selection
- Spatial distribution methods for stationary sources
- Temporal distribution methods for stationary sources
- Projection methods for stationary sources
- Spatial and temporal distribution and projection methods for motor vehicles
- Hydrocarbon species allocation.

In addition, detailed example descriptions of inventory preparation methodologies used in the San Francisco Bay Area and in the Denver metropolitan area are contained in Appendix C.

Grid Selection

A primary planning consideration for photochemical modeling is the determination of the grid system. It is extremely important to develop an appropriate grid system at the start so that all the source emission estimates are referenced to one grid system.

The area covered by the emission grid should be large enough to 1) include all major emission sources in the region within the grid, 2) include the receptor areas where highest pollutant concentrations are expected, 3) encompass areas of future industrial and residential growth, 4) include as many ambient pollutant monitoring stations as possible for model performance evaluation, and 5) include key features of the prevailing meteorology in the region.

Conversely, the size of the individual grid cells should be as small as possible in order to maximize the spatial resolution of the impact of individual sources. However, smaller grid-cell size requires additional effort to collect the necessary emission data and to develop appropriate techniques to allocate source emissions into the grid cell level.

A standard grid system utilized as the coordinate system for photochemical modeling is the Universal Transverse Mercator (UTM) system. This system is used in the NEDS and EIS emission data systems to reference all point source locations. Other common grid systems are the latitude and longitude system, and the State Plane Systems. However, it is very important to use a uniform coordinate system, because changing from one coordinate to another is not straightforward. For example, changing from state plane coordinates to UTM coordinates is a difficult procedure involving computerized calculations.

A typical grid size in a number of urban areas where photochemical models have been run is about 80 km by 80 km, with grid cells typically being about 2 x 2 km. A typical LIRAQ grid is shown in Figure 3-1 for the San Francisco Bay Region; the grid is 100 km by 100 km and consists of 5 x 5 km grid cells.

Overview of the Inventory Effort

There are three different types of inventories that may be prepared for control strategy planning purposes:

- o Prototype day inventory - This is the inventory applicable to an average day in the historical year that includes the prototype meteorological days being modeled, and is to be used in the model performance evaluation. Ideally, the inventory should correspond to the prototype day(s) being modeled; however, in practice tailoring the inventory for specific days is possible for very few source categories (e.g., electric utility boilers). Seasonal adjustment factors may be used.

- o Baseline projection inventories - These are inventories appropriate to an average day in future years that are projected assuming that no new controls are implemented beyond those already adopted. For example, the effects of continued implementation of the federal motor vehicle control program should be included in these inventories as estimated in EPA's MOBILE1 computer code. Appropriate future years for baseline inventories might be 1982 and/or 1987 to address attainment requirements of the Clean Air Act, and possibly 2000 to address long term maintenance requirements.

- o Control strategy inventories - For each alternative control strategy to be tested, a separate inventory should be produced that includes the estimated effects of the strategy on emissions in each future year of interest.

Reproduced from best available copy.

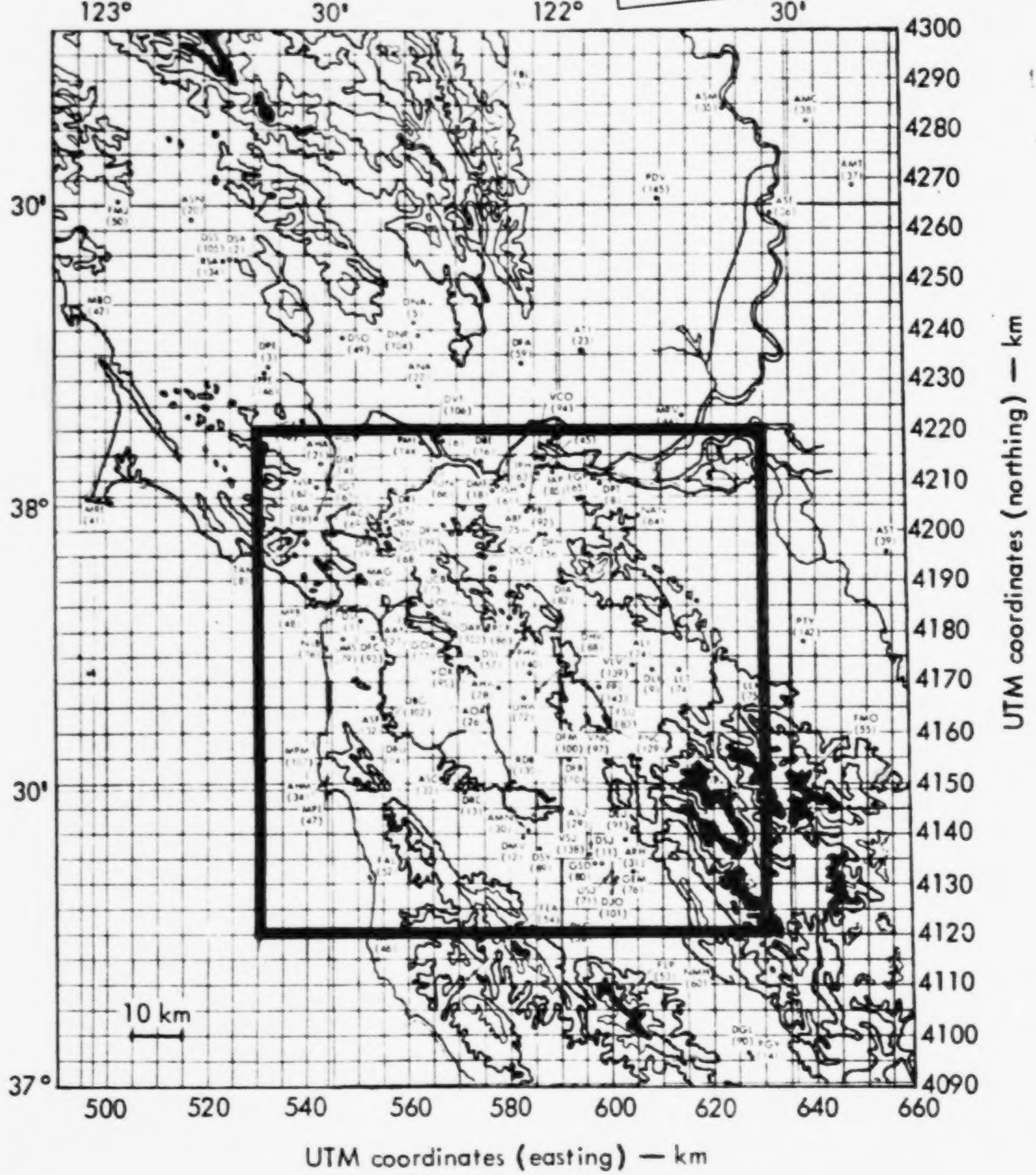

123°
30'
122°
30'
30'
38°
30'
37°
4300
4290
4280
4270
4260
4250
4240
4230
4220
4210
4200
4190
4180
4170
4160
4150
4140
4130
4120
4110
4100
4090
500
520
540
560
580
600
620
640
660
10 km
UTM coordinates (easting) — km
UTM coordinates (northing) — km

Figure 3-1. Map of Bay Area showing the region within which 5-km grid resolution is available in LIRAQ.

The prototype day inventory is needed to check model performance against ambient monitoring data. The baseline inventories are useful for determining what future air quality will be and whether additional control programs beyond those already adopted are necessary to meet the ozone standard. The control strategy inventories are useful for simulating the effects of alternative strategies and whether they will result in attainment and/or maintenance of the ambient ozone standard.

Each of these inventories must be disaggregated spatially into grids, temporally (usually by hour of the day), and by species (usually into two or more hydrocarbon classes and NO and NO2).

The preparation of these inventories requires an investment of substantial manpower and computer resources above and beyond that normally devoted to the preparation of emission inventories. Estimates of costs to produce a complete set of inventories may vary from $20,000 to $250,000 depending on the size of the region, the completeness of existing data bases and computer codes in place, and the skills of available staff.

A final note before proceeding into more detailed discussions is that as all of these inventories are being developed, care must be taken to ensure consistency from one inventory to the next. The methods used to disaggregate the inventories should be the same for prototype day, baseline, and control strategy inventories. Energy assumptions, growth assumptions, etc. should also be consistent. In this way, the differences from one inventory to the next will in fact reflect the changes that are desired to be measured and evaluated through the modeling process, rather than artificial differences introduced through the use of inconsistent methodologies.

Spatial Distribution of Stationary Source Emissions

For the purposes of this document, stationary sources are classified as point, airport, and area sources. Point sources are generally considered to be those that emit at least 100 tons per year of air pollutant from a stack or group of stacks. Area sources are individually small emissions that are impractical to consider as separate point or line sources. It should be noted that the value 100 tons per year is somewhat arbitrary, but it is in common use; smaller cutoff levels for point sources would improve the inventory for modeling purposes. Airports are included as a separate category because emissions from airports are neither spread over an entire county nor concentrated in a single stack.

Point Sources. Point source data are available nationally from the NEDS and EIS emission systems and locally from air pollution control agencies in the modeling region. Regardless of the data source to be used, it should be specifically reviewed to determine whether the data for each source is up to date and conforms with the input requirements of the model to be used.

In general the major point sources present no serious problems in spatial distribution. One can pinpoint their locations with a street map, and read their UTM coordinates with a USGS map to the nearest 0.1 km. Major point sources are carried as separate listings and may be further divided according to stack height. The stack height categories differ from one model to another. The SAI model injects the emissions into an appropriate layer based on the estimated plume rise for each individual point source while LIRAQ treats only two categories, "surface" and "elevated" categories, depending on whether the stack height is greater or less than 100 feet (30.5 meters).

For preparing photochemical modeling inventories the following information for point sources is very useful and should be included in point source listings.

- Location of UTM grid coordinates
- Source type
- Annual emission rates for each pollutant
- Normal operating hours
- Seasonal distribution of emissions

For models that have vertical resolution of pollutants the point source inventory should include stack parameters: stack height, stack diameter, exit gas temperature, and gas flow rate.

Airports. Since airports also have known locations, some inventories treat airports as point sources while others treat them as an independent source category as a matter of convenience. Because some landing and take-off emissions are spread out within the mixing layer, it is more appropriate that airport emissions are distributed over neighboring grid squares in addition to the several grid squares where the airport is located.

Area Sources. For air quality modeling, aggregated area source emissions must be distributed in some fashion over the geographical region of interest.

The simplest approach would be to distribute the emissions uniformly over a given study area, but this approach might result in a substantial fraction of area source emissions being distributed over undeveloped areas (e.g., mountain ridges, bay, marshlands, and undevelopable lands, etc.). The next level of sophistication would be to distribute the emissions according to population. This would be a great improvement, but some major flaws remain. Most census data concern residential population only and would misplace the many non-major point sources which operate in industrial and commercial areas. A more accurate distribution can be achieved, however, if source activities can be correlated with some sub-group of population, or with categories of employment or land use.

Due to lack of appropriate procedures and data, a commonly used spatial distribution method for area sources in many modeling studies is the spatial distribution based on population or uniformly gridded emission factors for area sources. In most metropolitan areas, local agencies have some sort of land use data. Land use information is normally available by land use categories (e.g., residential, commercial and industrial, etc.) and by certain planning zones. From these land use data, one can set up an allocation procedure that involves determining the relationship of each type of land use to each emission source category.

A cross classification of source categories with land use data can be fairly time consuming, depending upon the detail of the land use data available in a given region. However, the use of surrogate indicators derived from land use categories is often the best available procedure for apportioning regional area source emissions to individual grid cells. (See Perardi, et al., 1979, for more details.)

Temporal Distribution of Stationary Source Emissions

In order for photochemical simulation models to adequately predict hourly ozone and other pollutant concentrations, typical hour-by-hour emissions estimates are needed at the grid cell level. Several basic approaches can be used for providing the temporal detail needed in the inventory used in a photochemical model. The most accurate and exacting approach would be to determine the emissions for specific sources for each hour of the time period being modeled. This approach is generally most applicable for point sources. An alternate approach is to develop typical hourly patterns of activity levels for each source category and then apply these to the annual or seasonally adjusted emissions to established hourly emissions. This approach is more appropriate for area sources.

Normally, the photochemical air quality model, and therefore the emission inventory, is applied during the season of the year in which weather is most conducive to ozone formation; for most locations, this means the summer months. Similarly, emissions are usually chosen to represent the day of the week on which pollutant-emitting activities are at a maximum, normally an average weekday. In this case the prototype days would be summer weekdays. Seasonal adjustments should be made first, to change the aggregate source inventory from an annual average basis to a summer weekday, typical of the ozone season.

Airports. Aircraft emissions generally are divided into three inventory classifications: commercial carriers, military, and general aviation. Commercial carriers are most important from an emission viewpoint, and fortunately these are also the best documented. Comprehensive schedule books keep up-to-date listings of commercial flights to all the major airports of the world with arrival and departure times and aircraft type. Temporal resolution factors for commercial airports in a given area could be compiled based on this kind of data.

Area Sources. The hourly distribution of area sources can be based on diurnal variation coefficients (percent of total daily emissions emitted in each hour) for each source classification. These coefficients could be compiled by source inventory engineers familiar with the types of sources in each classification.

Stationary Source Projection Methods

Because of the need to determine whether a given area will achieve the ozone standard, or remain in compliance in future years, the inventorying agency must develop emission projections for future years for use in the photochemical models.

There are many approaches for making emission projections; perhaps the simplest and most frequently used technique is extrapolation or regression where a series of historical data points are projected into the future in some simple linear fashion. Here the historical emissions estimated are correlated to any number of variables (e.g., population, fuel use, etc.). What results is an equation which represents the "best" correlation of the variables to changes in emission rates for use in projecting future emissions. In essence, it is assumed that the factors which have apparently accounted for historical emissions will continue to do so in the future in the same manner. To forecast emissions using this technique requires a forecast of all the variables assumed to be important in accounting for historical emission trends.

Another emission projection approach is the use of surrogate variables. This procedure assumes that a given pollution source category can be accurately projected by projecting some related variable. For example, increases in aircraft emissions might be based on projected increases in passenger air travel. Similarly, hydrocarbon emissions from internal combustion engines might be projected on the basis of forecasts of population growth.

Frequently, very little data is available to assist in making future projections. In these cases, engineering or scientific judgment can be used. (See Leong and Wada, 1978.)

There are two EPA publications which deal with methodologies and procedures for estimating emissions in the future years: the Regional Emission Projection System (REPS) (see Booz-Allen and Hamilton, 1974), and the Guideline for Air Quality Maintenance Planning and Analysis, Volume 7: Projecting County Emissions (U.S. EPA, 1975). In general, the methodologies and data bases employed or recommended in these documents provide very crude estimates of projected emissions, which should be carefully reviewed before being used in a photochemical model application.

The basic emission projection methodology employed in the San Francisco Bay Area is the use of surrogate variables. For each of 107 source categories, related variables are defined and necessary data for projecting related variables are studied. This procedure assumes that a given pollution source category can be accurately projected by projecting some related variable or variables. For example, future emissions from aircraft in the Bay Area were estimated from projected air passenger travel, aircraft type and emission factors in a given future year. Table 3-2 presents an example of some source categories and related variables used for projection.

TABLE 3-2

TYPICAL VARIABLES FOR EMISSION PROJECTION FOR THE BAY AREA

SOURCE CATEGORY	VARIABLES USED FOR PROJECTION
Petroleum Refining	Oil supply
Chemical Processing	Oil supply and population
Industrial Coating	Population
Petroleum Evaporation	Oil supply and gasoline demand
Other Organic Compound	Population and other factors
Dry Cleaning	National trend
Domestic Combustion	Population
Industrial Combustion	Natural gas and other factors
Power Plant	Fuel usage, natural gas usage and coal usage
Aircraft	Aircraft operations and air passenger miles

Motor Vehicle Emissions

Inventory guidelines previously referenced have described alternative methods for preparing disaggregated mobile source inventories. The method best suited for planning purposes involves the use of previously developed computerized transportation models. Most states and/or major metropolitan areas use such models to assist their transportation planning and decision-making processes. They have the capability of simulating the effects of a variety of (though not necessarily all) transportation controls. Use of these models in cooperation with transportation planning agencies will ensure proper projection and spatial and temporal disaggregation of motor vehicle emissions.

The standard system used nationwide for forecasting regional travel patterns and highway and transit needs consists of a series of models as follows:

- o Trip Generation Model
- o Trip Distribution Model
- o Modal Split Model
- o Network Assignment Model

Two key outputs from these models for air quality evaluation are: (1) a trip table indicating the number of daily trips originating and ending in each traffic assignment zone*; and (2) an "historical record" file that indicates average daily traffic volumes and average speeds on a geographically coded highway network.

A number of steps are required to convert the output from the transportation models into gridded, hourly emission estimates suitable for input to a photochemical model:

- o Appropriate motor vehicle emission factors must be obtained.
- o There must be a computer code that reads the transportation data and applies the appropriate emission factor.
- o The resulting emissions must be assigned to the appropriate grid, and disaggregated to each hour of the day, and possibly by species depending on the photochemical model to be used.
- o The data must be arranged in the appropriate format for input to the photochemical model.

---

*Traffic assignment zones are cells of irregular size and shape which subdivide the urban area or study region to the degree of detail necessary for transportation modeling.

A decision should be made as to whether trip-end emissions (emissions due to cold starts, hot starts, and hot soaks) should be treated separately or lumped with hot-stabilized emissions. It is evident that the spatial distribution of trip-end emissions would be somewhat different from hot-stabilized emissions, although whether the differences are significant enough to merit separate consideration is unknown at this time.

Figure 3-2 illustrates the process for preparing separate inventories for trip end and hot stabilized emissions. EPA's MOBILE1 code (U.S. EPA, 1978) for computing motor vehicle emission factors is exercised a number of times to produce separate hot stabilized and trip end emission factors for a given future year at different average speeds and ambient temperatures. Hot stabilized emission factors are then applied to the historical record file output from the transportation models to compute average daily emissions on each highway segment. These data are then disaggregated into hourly portions and assigned to the appropriate grid squares covering the region to be modeled. The standard Federal Highway Administration codes SAPOLLUT and SAPLSM (U.S. DOT, Federal Highway Administration, 1976) may, with only slight modification of the emission factors, be used to compute emissions and assign them to grids as stated.

Trip end emission factors are applied to the trip table, and the resulting emissions are assigned to traffic assignment zones. To convert these data into gridded emissions, a method for overlaying a regular square grid onto the system of irregularly shaped traffic zones must be developed. In its most straightforward form, this overlay would consist of a conversion table that lists the percentage of each traffic zone lying within each grid. (This table may also be used to allocate intrazonal trip emissions to grids.) Other methods are possible, and since there are no standard programs for preparing gridded inventories from data on a traffic zone basis, the model user is left to exercise judgment and ingenuity in the preparation of the inventory.

After both hot-stabilized and trip end emission data files are prepared, they may be merged into a single motor vehicle emissions file.

The simpler alternative of using composite trip end and hot stabilized emission factors eliminates the use of the trip table as well as the need for a method to convert traffic zone information to grids. (As long as intrazonal trip emissions are a small component, they may be represented by adding small fictitious links to the network.) In Figure 3-2, this alternative would consist solely of the left half of the flow chart using composite emission factors rather than hot stabilized emission factors.

Hydrocarbon Species Allocation

Because air quality models attempt to simulate complex photochemistry, hydrocarbon emissions must be distributed into various species classes. In addition, NOx emissions have to be distributed into NO and NO2.

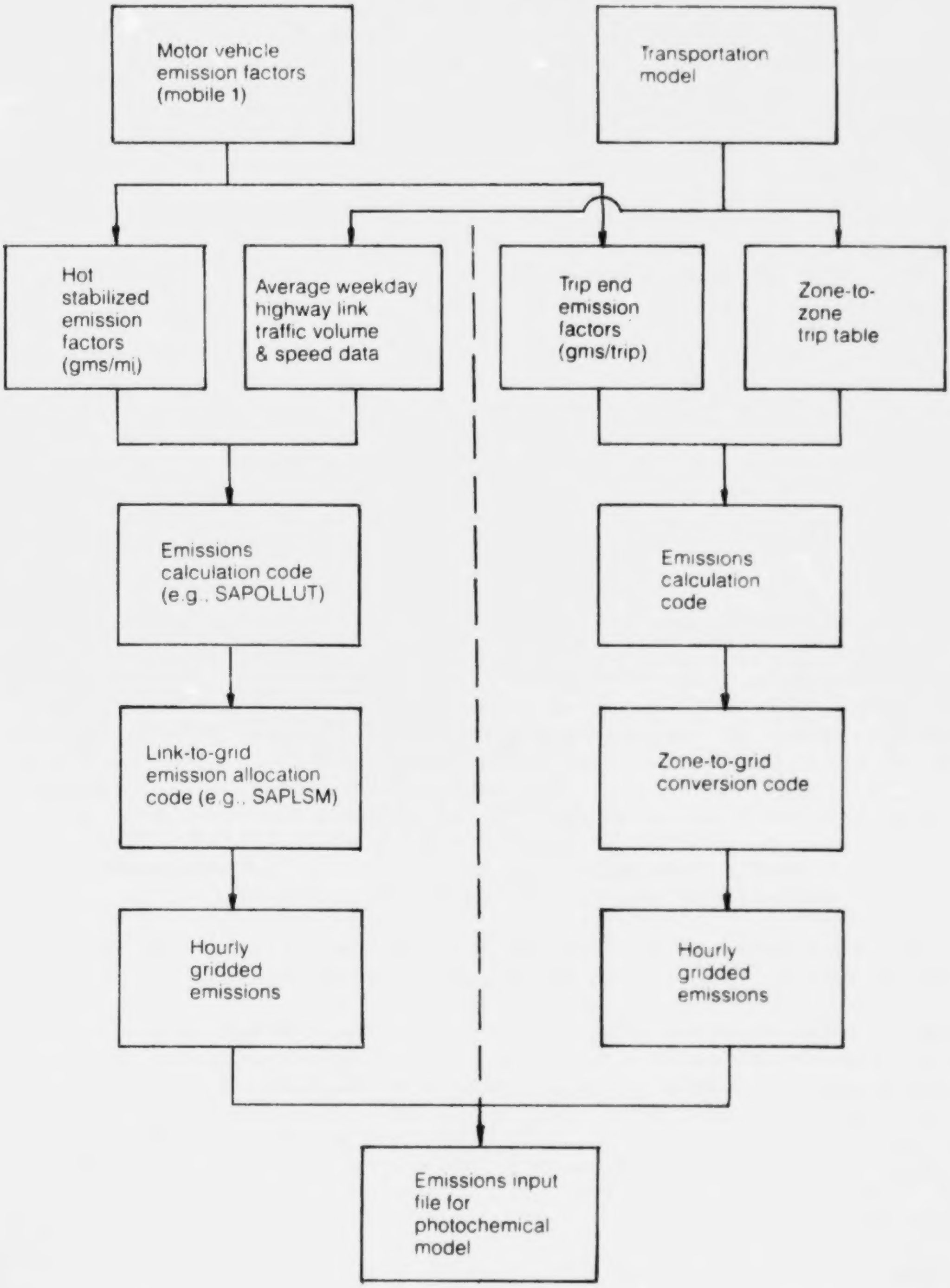

Motor vehicle emission factors (mobile 1)
Transportation model
Hot stabilized emission factors (gms/mi)
Average weekday highway link traffic volume & speed data
Trip end emission factors (gms/trip)
Zone-to-zone trip table
Emissions calculation code (e.g., SAPOLLUT)
Emissions calculation code
Link-to-grid emission allocation code (e.g., SAPLSM)
Zone-to-grid conversion code
Hourly gridded emissions
Hourly gridded emissions
Emissions input file for photochemical model

Figure 3-2

Process for preparation of separate, disaggregated trip end and hot stabilized motor vehicle emission inventories

The typical procedure for allocating hydrocarbons to species classes is (1) to assume that the total hydrocarbon emissions from each type of source contain a certain percentage of each class of compound, and (2) to apportion the total hydrocarbon emissions by these percentages and thus determine emission totals of each species class by source category.

Each photochemical model requires specific types of volatile organic compound (VOC) input data from an emission inventory, and this VOC information must be in proper format for use by the model in this computations of chemical reactions. The EPA report, "Volatile Organic Compound (VOC) Species Data Manual," by KVB Engineering, Inc. (EPA-450/3-78-119) may provide useful data in this regard.

For example, the LIRAQ model contains a 48-reaction chemical mechanism with three main classes of hydrocarbons, designated HC1 (mainly alkene), HC2 (mainly alkanes, simple aromatics, ethers, alcohols), and HC4 (mainly aldehydes, some ketones and some aromatics). Table 3-3 summarizes the percentage breakdown of hydrocarbon into the three LIRAQ classifications by source type as used in the San Francisco Bay Area. Such estimates should in general be made in consultation with personnel familiar with the composition of organic emissions and the treatment of atmospheric chemistry in the model of interest.

AMBIENT AIR QUALITY DATA

This section discusses the basic air quality monitoring data that are needed to set initial and boundary conditions and evaluate the performance of photochemical models. The discussion is followed by two examples of efforts used to gather data for performance evaluation.

The evaluation of a photochemical model's performance primarily consists of comparing the ambient pollutant levels predicted by the model with the actual ambient pollutant levels measured by a monitoring network. The purpose of this comparison is to determine the accuracy of the model predictions. Theoretically, the most accurate way to achieve this comparison is to insure that the spatial and temporal characteristics of the predicted and measured ambient concentrations are similar. For example, if the model output is expressed in terms of 1-hour ozone concentrations averaged over a grid area of 1 square kilometer, then the ambient data should match this as closely as possible.

The following ambient air quality data and related information are needed to evaluate the performance of a photochemical model:

- measured hourly average ambient concentrations of HC, NO, NO2, and O3 (and in some cases CO and SO2);
- the location (geographical coordinates and altitude) of each monitoring station at which the measurements were made;
- the analytical methods that were used to measure the ambient levels of each pollutant.

TABLE 3-3

SUMMARY OF LIRAQ COMPATIBLE PERCENT EMISSIONS BY SOURCE TYPE

Source Type	HC1	HC2	HC4
Petroleum Production	-	100%	-
Refinery Operations	23%	77%	-
Underground Storage	15%	85%	-
Auto Filling Tanks	27%	73%	-
Fuel Combustion	38%	30%	32%
Waste Burning Fires	22%	40%	38%
Heat Treated Coatings	20%	80%	-
Air Dried Coatings	14.4%	85%	0.6%
Petroleum Based Dry Cleaning	2%	98%	-
Dry Cleaning - syn	-	100%	-
Degreasing TCE	100%	-	-
Degreasing III-T	-	100%	-
Rotogravure Printing	7%	93%	-
Flexigraphic Printing	21%	78%	1%
Rubber, Plastic, etc. Mfg	36%	51.5%	12.5%
Pharmaceuticals	27%	70%	3%
Misc. Solvents	17%	80%	3%
Gasoline Exhaust	22%	71%	7%
Gasoline Evap (Mobile)	24%	76%	-
Diesel Exhaust	18%	60%	22%
Gas Turbine (Jet)	33%	53%	14%
Piston Aircraft	26%	66%	8%

For any given area, the above data usually can be obtained from a number of sources:

- local air pollution control district;
- state air quality agency;
- Federal monitoring programs;
- universities and colleges;
- private industry;
- private research organizations.

The maximum amount of ambient air quality data that are compatible with model output should be obtained from each source.

Once the initial data-gathering exercise has been completed, the data should be screened prior to use in evaluating model performance.

The purpose of this screening is to ensure that:

- pollutants measured by different networks (e.g., local, state, Federal, private, etc.) were measured with comparable analytical methods;

- each station has enough data to obtain a statistically representative sample of the true pollutant levels in the air.

The screened data quite possibly may need to be adjusted before they can be directly compared to model output. For example, many monitoring systems measure total oxidant (of which ozone comprises a large fraction), while many models predict only ozone. The total oxidant air quality data would thus have to be corrected to represent only ozone. Furthermore, monitoring systems usually measure total hydrocarbons (THC) and non-methane hydrocarbons (NMHC), whereas photochemical models predict hydrocarbon levels by reactivity class. Some correction factor would have to be applied to the ambient hydrocarbon data to allow comparison with the model results. Even when prepared with the utmost care, such comparisons are difficult due to the high uncertainty in the monitoring data for NMHC.

The air quality data should next be examined to determine if they are adequate to verify the model. Specifically, the following issues should be considered:

- is the spatial and temporal resolution of the ambient data adequate to characterize initial conditions and to simulate the diurnal cycle in oxidant production?

- are data available to determine vertical concentration profiles of pollutants, and to determine boundary inflow through the top of the mixed layer?

- are there adequate data collected at stations outside the model domain to characterize boundary inflow?

If the existing data are inadequate to evaluate model performance and/or specify model inputs, then field studies should be employed to gather the needed air quality data. These field studies should also measure meteorological parameters concurrently with the ambient air pollutant concentrations.

The following paragraphs discuss two examples of collecting data to be used in evaluating the performance of photochemical models. The first example concerns the data collection effort used to evaluate the performance of LIRAQ in the San Francisco Bay Area, and the second example illustrates the monitoring that was undertaken to evaluate the performance of the Denver Air Quality Model (early version of the SAI Model) for Denver.

The data gathering effort for evaluating the performance of LIRAQ was aimed at collecting ambient O3, HC, NO and NO2 monitoring data from the BAAQMD monitoring network, a NASA aircraft monitoring program, and Barringer spectrometry studies. (These are described in greater detail later in this section.) These data were screened to ensure that they represented measurements taken with the same (or similar) analysis methods and that each station had enough data to formulate representative statistics. All of these data represented point measurements, and thus were not strictly comparable with the LIRAQ output, which is in terms of average concentration in a 1- by 1-, 2- by 2-, or 5- by 5- kilometer box (MacCracken, et al., 1975); however, since better data were not available, the point measurements were used to evaluate the model. (The problem of comparing point measurements with spatially-averaged model predictions is common to all photochemical models currently in use.)

After screening, the ambient hydrocarbon and oxidant data had to be adjusted to facilitate comparison with LIRAQ output. The ambient hydrocarbon levels predicted by LIRAQ were broken down into three reactivity classes, but the ambient levels measured by the monitoring stations were reported as either total hydrocarbons or non-methane hydrocarbons. The model output was modified slightly so that the predicted ambient hydrocarbon levels were reported in terms of a "typical" hydrocarbon type that could be compared with the ambient measurements. The pre-1975 measured ambient oxidant levels were artificially high due to inadequacies in the instrument calibration procedures used at the time. All the affected oxidant data were multiplied by 0.8 to correct the error (MacCracken, et al., 1975). The ozone levels predicted by the model were compared with the adjusted oxidant data.

Even though the BAAQMD (and other groups) had compiled an extensive ambient air quality data base, there was still insufficient data to develop and evaluate the performance of LIRAQ; consequently, field studies were undertaken to gather the needed data.* Specifically, NASA aircraft and Barringer correlation spectrometer field studies were performed to gather additional data to supplement the existing BAAQMD monitoring network. The NASA aircraft data consisted of measured ambient levels of CO, oxidant, O3, NO, NO2, and SO2, in addition to concurrent measurements of several meteorological parameters. The data provided vertical profiles of pollutant levels over both urban and rural sites (MacCracken, et al., 1975). Although these data were point measurements (and hence not strictly comparable with LIRAQ output), they did provide valuable information on the vertical profile of the pollutants. This information was used to evaluate the performance of LIRAQ in generating vertical pollutant profiles. The Barringer Correlation Spectrometer Study was conducted in order to obtain spatially-averaged ambient pollutant concentration data that could then be compared with the spatially-averaged LIRAQ output. Using mobile monitors, the total NO2 level (ppm) in a straight line path was measured in combination with fixed surface monitors measuring total sulfur and total NOx (MacCracken, et al., 1975).

The Denver Study (Anderson, et al., 1977) conducted during 1975-1976 contained another example of supplemental monitoring that was undertaken to obtain data for model evaluation. The spatial coverage of the ambient air quality data collected by the existing monitoring network was inadequate to evaluate the performance of the model.

The data used to evaluate the performance of the DAQM came from four sources:

- the existing ambient air quality monitoring network operated by the Air Pollution Control Division (APCD) of the Colorado Department of Health;
- supplemental ambient air quality monitoring stations operated by the Colorado Highway Department during 1975 and 1976;
- the meteorological monitoring network operated by the Colorado APCD;
- supplemental meteorological monitoring stations operated by the Colorado Highway Department, the National Weather Service, private concerns, and others.

Each of these data sources will now be discussed.

---

*Much of the field work was originally undertaken to gather data for model development; this data was later used to evaluate the performance of LIRAQ.

The existing Colorado APCD ambient air quality monitoring network consisted of six stations measuring ambient HC, NO, NO2, O3, SO2 and CO levels; only HC was monitored at all six stations during the time of the study (1975-1976).

The supplemental ambient air quality monitoring network consisted of four stations in 1975 and three stations in 1976. Only O3 was monitored at the supplemental stations. The supplemental network consisted of one mobile and three fixed stations in 1975 and one mobile and two fixed stations in 1976. In 1975, the supplemental monitoring was done on predicted "bad" ozone days; approximately 20 days of data were collected. In 1976, the monitoring system was run continuously for 90 days during the peak ozone season.

The meteorological monitoring was conducted at four of the six ambient air quality stations. Wind speed and direction at 10 meters above the ground were monitored at the four stations.

The supplemental meteorological monitoring was conducted by a variety of sources. The exact number of contributing stations varied daily, with a maximum of about 11 stations supplying data. No meteorological variables were monitored at the supplemental ambient air quality stations.

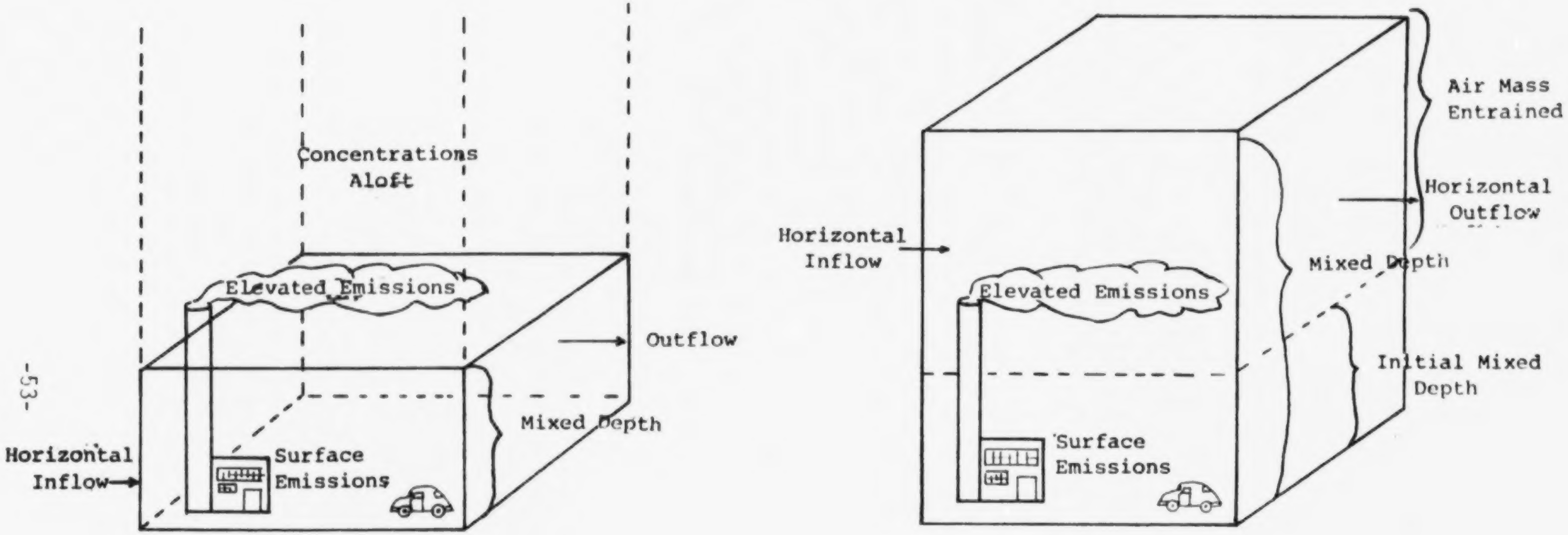
TREATMENT OF INITIAL AND BOUNDARY CONDITIONS

The treatment of initial and boundary conditions can play a major role in determining the results of an air quality model simulation, especially for ozone/oxidant. Primary pollutants are introduced to the grid model domain* through three model elements of comparable magnitudes: initial conditions, inflow through the boundary, and emissions. The actual and relative magnitudes of these three elements will be a function of detailed model inputs (e.g., boundary concentration, emissions rate) of location within the model domain, and of model structure. For example, Lagrangian models generally assume no flow across horizontal boundaries, but are influenced by initial conditions and entrainment at the top of the air column. Because individual trajectories are often run for only a few hours, the relative importance of initial conditions usually be greater for a Lagrangian model than for an Eulerian model.

Initial conditions completely determine the pollutant content of the model at the initial time, and dominate the loading of pollutants for some time thereafter. Figure 3-3 illustrates a single element on an upwind boundary of an Eulerian grid model. The initial mass loading of pollutant in a given model element (assuming uniform concentration over the cell) will be given by the product of the initial mass concentration in the element and the initial volume of the element.

---

*The volume of space to which a model is being applied, usually defined horizontally by the extent of the grid, and vertically by the ground and the height of the mixed layer.


Concentrations Aloft
Elevated Emissions
Outflow
Mixed Depth
Horizontal Inflow
Surface Emissions
Air Mass Entrained
Horizontal Outflow
Horizontal Inflow
Mixed Depth
Elevated Emissions
Initial Mixed Depth
Surface Emissions

a. A representative grid cell on the upwind boundary at the time of problem initialization in the early morning.

b. The same grid cell at the time of maximum ozone. The air mass between the initial time and the time of the ozone peak will have entered through one of the boundaries. It will either reflect entrainment of air from above the mixed layer or from convergent flow.

Figure 3-3. Schematic Diagram of Grid Cell Inputs

Inflow across the model boundaries will provide a pollutant source given by the product of the volume influx times the concentration at the boundary integrated over the time period simulated. Inflow often dominates the pollutant mass input near the inflow boundaries.

Emissions provide a source of pollutants given by the integral of the emissions rate over the time period simulated. Emissions will usually govern mass input near strong sources, but may not be the primary source of pollutant a few kilometers away from major sources.

The relative importance of these three terms will vary greatly as a function of time and location. Emissions will usually dominate the mass inputs of primary pollutants in central cities; however, this will often not be the case for the modeling grid if it is structured to consider ranges comparable to those involved in the formation and transport of photochemical oxidants.

In a representative 0400 hours to 1800 hours LIRAQ simulation of Bay Area air quality, emissions accounted for about 3/5, initial conditions accounted for about 1/10, and boundary influx accounted for about 3/10 of the total input of hydrocarbons and NOx to the model domain. For CO, emissions accounted for about 1/5, initialization for 1/10, and boundary flux for 7/10 of the total mass input. Although these numbers would change if the meteorological scenario, grid domain, model structure or assumed boundary and initial concentrations had been changed, they are representative of a smoggy San Francisco Bay Area day. Other areas and other models might receive relatively more mass input from initial conditions and less from boundary flow, since early morning mixing depths are often very shallow in the San Francisco Bay Area, and a strong sea breeze is often observed even on smoggy days, giving rise to a significant boundary flow.

Just as initial conditions completely determine concentrations at the beginning of a model simulation run, boundary fluxes will often determine the mass loading near the upwind boundaries of Eulerian models. This happens because the upwind boundaries are often in regions of low to moderate emissions, and even a modest inflow velocity can carry a large volume of air through a grid cell boundary.

It might seem that this discussion overstates the importance of a few ppb of pollutants introduced as background values on initialization or through boundary fluxes. Indeed, relative to ground level measurements of primary pollutants in central urban areas, background levels are usually unimportant. There are exceptions; for example, for more than half of a 100 km x 100 km grid centered on the city of St. Louis in a recent LIRAQ calculation (at 1500 hours on a high oxidant day), mean layer CO was within 15% of the background value used (100 ppb), the mean layer hydrocarbons were less than 3 times the assumed background of 30 ppb, and mean layer NOx was less than twice the asssumed background of 4 ppb. The case is qualitatively similar in the San Francisco Bay Area, especially in suburban high oxidant areas. Although concentrations of hydrocarbon and NOx are not particularly high in many suburban locations, the influence of moderately low concentrations of hydrocarbons and NOx on ozone is still significant, at least in the LIRAQ model. This was also established in sensitivity studies carried out with the 1975 version of the LIRAQ model and described in the LIRAQ documentation (Duewer et al., 1978). These studies found that the difference between using very small boundary concentrations and boundary conditions slightly larger than those now employed was a factor of nearly 1.6 in the average ozone content at measuring stations (a larger factor if averaged over the entire domain).

Methods for Treating Initial Conditions

Base Year

Ideally, in treating base year scenarios the initial conditions are based on measured ambient concentrations with the measurement stations giving a fairly uniform coverage of the entire model domain, and providing extensive data aloft. In practice, base-year initial conditions are usually based on relatively sparse measurements taken almost exclusively at the surface and primarily near urban centers. As a result, concentrations aloft and away from urban centers must be estimated from the limited available direct measurements, from conjectures about concentrations at the domain boundary and aloft, and from some interpolation/extrapolation scheme. This is largely a consequence of the great difficulty and expense involved in accurately measuring pollutants at concentrations one to two orders of magnitude below standards, or of measuring anything aloft. An adequate sampling program would be expected to cost hundreds of thousands of dollars if run for even a few days.

In LIRAQ applications, a vertical profile based on emission rate, deposition velocity, mixed layer depth, and wind speed is combined with surface observations to yield mean layer concentrations above measurement stations. Concentrations at boundaries are estimated based on the larger of global background or lowest measured value. If data had been available that indicated higher concentrations, higher concentrations would have been adopted, but the available data were of the character of upper limits. Currently in LIRAQ applications, a Gaussian weighting scheme is used to interpolate initial concentrations over the grid; the interpolation uses measured values and boundary conditions at the horizontal boundaries. Finally, concentrations in grid cells that are above the inversion are set to the values estimated at the upper boundary.

In the case of the SAI Urban Airshed Model a similar procedure has been used except that it is a multi-layer model. The mixed layer is assumed to be well mixed, the concentrations above the mixed layer have been set to an assumed background value (or available measurements aloft, if any, have been used), and the concentrations assumed at the horizontal boundaries have not been used. In order to avoid extrapolation of urban center concentrations into rural areas, synthetic stations are occasionally inserted into outlying areas and assigned background concentrations. In areas (such as Los Angeles) where measurements show high concentrations of pollutants aloft or at the boundaries, high concentrations have been used in modeling efforts.

In the case of LIRAQ applications, initial concentrations usually provide on the order of 10% of the total input of most primary pollutants for a typical 14-18 hour run. For some trajectories in trajectory model calculations, initial conditions will provide nearly all of the input of primary pollutants. This will be particularly true if the trajectories are run for short periods. In application of a multi-layer model, if the maximum height of the model were to be maintained at a constant value, and the mixing height varied within that, initialization would occur over a much deeper layer, and initial conditions would usually provide a larger fraction of the total mass input to the model. Of course, boundary flow would be correspondingly reduced in importance.

Since the photochemical models use multiple classes of reactive hydrocarbons, measured or estimated nonmethane hydrocarbon concentrations (NMHC) must be converted into alkene-like, alkane-like, aldehyde-like, etc. hydrocarbon concentrations. The choices made here rarely have any directly appropriate data to guide them and can have factor-of-ten effects on the effective reactivity of the hydrocarbons entering the model through initial conditions. The factors needed to assign NMHC to model classes should be strongly affected by the local emissions inventory, the significance of biogenic emissions (which should provide a few ppb of very high reactivity materials) and long range transport (which should provide primarily low reactivity materials). The breakdown of NMHC into model hydrocarbon classes can be expected to differ from model to model, from region to region, and even from day to day for a given region. The ideal resolution of this problem would require a comprehensive and detailed analysis of the actual organic species contained in ambient air samples in the region of interest, preferably on the days to be modeled. However, a significant improvement in the available data is definitely feasible for a cost likely to be of the order of ten thousand to fifty thousand dollars.

Future Year Simulations

For future year simulations, no measured initial conditions are available, and initial concentrations must be forecast from emissions changes, baseline observations, and known background concentrations. In cases where the measured base year initial concentrations are substantially above background concentrations, a reasonable hypothesis is that the measured concentrations reflect controllable local emissions. However, if the base-year initial concentrations approximate global background concentrations, future year initial concentrations will not likely be much different from base-year initial concentrations (unless emissions are to be dramatically increased).

Because initial concentrations measured in the San Francisco Bay Area have generally been well above global background, in LIRAQ applications the initial concentrations have been scaled with emissions. This has been done in one of two ways. In some cases, the initial concentration field developed by the model was multiplied by the ratio of future emissions averaged over the grid and day to emissions on the prototype day similarly averaged. In what may be a slightly superior alternate methodology the station observations have been scaled by the same ratio, the initial boundary concentrations left unchanged, and a new concentration field developed. In the first method, all initial conditions are scaled with emissions, even when they may be dominated by global background concentrations. The virtue of the second method is that initial conditions are not scaled when they are in regions where the initial conditions are determined by the assumed background values. In the application of the SAI model to Denver the initial concentrations were near background levels, and no scaling was used.

For both of these examples, the estimation of future initial concentrations of precursors may be expressed by the following formula:

$$C(\text{future}) = C(\text{bkgd}) + \frac{E(\text{future})}{E(\text{prototype day})} \times [C(\text{prototype day}) - C(\text{bkgd})]$$

If background concentrations are small relative to the observed initial concentrations on the prototype day being modeled, then future initial concentrations become scaled by the ratio of emissions. If initial concentrations are close to background levels on the prototype day, then little or no changes should occur for future year simulations. The situation becomes more complicated if the background contribution to the observed initial concentrations includes pollutants transported from a nearby area not included in the modeling grid.

One possible method of reducing the problem of specifying initial conditions would be to perform extended (i.e., multiday) simulations. Before this approach is adopted, one must establish that the model adopted can satisfactorily represent such a multiday period, and that the available data permits a satisfactory representation of flow patterns aloft as well as at the ground.

Boundary Conditions

Most models use the concentration computed just inside the outflow boundary as an outflow boundary condition, and outflow boundary conditions present no difficulty. The problem is in the treatment of concentrations aloft (when the inversion base is rising) and at the upwind boundaries. In the ideal case, observational data are available to provide information about the concentrations at the model boundaries. However, in practice, very few useful data are ever available. In part, this is a result of the difficulty involved in making measurements aloft, and the tendency for most air quality measurements to be made at locations where standards are expected to be violated. However, a more pervasive problem is that most instruments designed for air quality measurements are either improperly calibrated or insufficiently precise when concentrations are at the ppb level. Carbon monoxide provides a simple example of both faults. In many air quality monitoring programs, carbon monoxide concentrations are rounded to the nearest 1000 ppb. When carbon monoxide has been measured using systems carefully designed to operate at levels below 1 ppm, the lowest concentrations measured over the north Atlantic or north Pacific have been about 100-150 ppb. Such concentrations would contribute about 1/2 the total mass loading to a typical model simulation over a 10,000 km2 grid in an area like San Francisco. However, when the RAPS program in St. Louis adapted CO monitors for operation at sub-ppm levels, they frequently reported concentrations of 50 ppb, a default value recorded whenever the measurements were 50 ppb or less. This is almost certainly a reflection of improper calibration (R. A. Rasmussen, private communication). The situations for reactive hydrocarbons, NOx, and SO2 are generally worse than for CO since the concentrations are generally much lower, and in the case of hydrocarbons, corrections must be made for 1,300-1,700 ppb of methane.

In practice, boundary conditions have occasionally been defined based on air quality measurements taken near the boundaries of the modeling region. When this was done in the early 1970's, the boundary flows often almost completely dominated the problem. In more recent applications much lower concentrations have generally been used. However, even at concentrations representative of the marine boundary layer or the upper troposphere, CO boundary fluxes would usually be comparable to model emissions over a 100 km x 100 km grid. The situation with respect to non-methane hydrocarbons and SO2 is less clear because true background concentrations of non-methane hydrocarbons and SO2 are not well established. In the case of NOx there is evidence that in the marine boundary layer, and presumably in the upper troposphere, NOx concentrations are in the range of 0.01 to 0.15 ppb (Noxon, 1978; McFarland et al., 1978). At these levels NOx boundary flow is negligible relative to emissions. However, there is ample room to doubt whether concentrations that low would be seen in continental air over the United States although very low NOx concentrations might be appropriate as western boundary fluxes in the San Francisco Bay Area.

This discussion does not imply that extremely low boundary concentrations are necessarily correct. Either recirculation or long-range transport might lead to high concentrations of anthropogenic materials at model boundaries, and, at least occasionally, natural events may also be important (e.g., volcanoes can be major sources of SO2, forest fires major sources of particulate and organic materials, normal biogenic and geogenic organic emissions may also be significant in some areas). The point is that the difference between two sets of fairly low concentrations at the boundary can potentially have a significant effect on model predictions of secondary pollutants at downwind sites.

Examples of the Treatment of Boundary Conditions in Baseline Calculations and Model Evaluation Studies

In LIRAQ applications, the inflow boundary conditions have been treated with an expression of the form

$$C_b = (C_o^2 + \rho C_i^2)^{1/2}$$

where Cb is the inflow boundary concentration, Co is a minimum background concentration, p is a parameter between zero and one, and Ci is the concentration just inside the boundary. Co and p are specified for each of five boundaries (E, W, N, S, Top). The rationale for this expression is that if Ci is low, the adjacent areas just beyond the boundary would usually also have low pollutant concentrations reflecting background levels; however, if Ci is large with respect to background, the boundary region likely either has significant emissions, or has been influenced by air advected from an emissions-dominated area. Also, in the case of secondary pollutants, air just outside the boundary should have experienced similar photochemical conditions to that just inside. In either case, pollutant concentrations in air from just outside the domain boundary would more likely be above background if concentrations just inside the boundary are elevated. The precise form was chosen to give a continously differentiable function for convenience in the mathematical treatment of the model. In the earlier LIRAQ developmental efforts p values of 0.5 to 0.95 were usually used (Duewer, et al., 1978). This suffers from the disadvantage that a p of 0.95 causes the concentration in a region experiencing no sources except rapid advection from an upwind boundary to eventually reach 3.2xCo. In contrast, a p of 0.5 would only cause the concentration to reach 1.15 Co. In current applications of the LIRAQ model all p's are recommended to be set at 0.3 except p for ozone aloft, which has been set at 0.85. This reflects the observed behavior of ozone in the San Francisco Bay Area, where ozone often displays its maximum concentrations aloft.

The Co values must reflect local conditions. In the San Francisco Bay Area carbon monoxide observations are seldom reported as less than 1.0 ppm, even at rural stations, although some suburban stations report 1.0 ppm most of the time on the days that have been simulated by LIRAQ. For this reason Co was set to 1.0 ppm (1000 ppb) for carbon monoxide, although a Co of 100-200 ppb would be easier to justify. Hydrocarbons are divided into three classes for LIRAQ: HC1, a high reactivity class similar to alkenes; HC2, a low reactivity class similar to alkanes; and HC4, a photoreactive aldehyde-like class. Representative Co values used in LIRAQ applications would be 10 ppb for HC1, 50 ppb for HC2 and 1 ppb for HC4. Co's for NO and NO2 have often been set at 4 ppb, SO2 at 10 ppb, and ozone at 10 ppb in the boundary layer, 25 ppb aloft. Species other than those listed have been assigned very low Co's of 0.004 ppb.

The SAI model has a substantial variety of possible methods for generating boundary conditions. In most applications, boundary conditions have either been based on measurements or have reflected "background" values. When the boundary conditions are based on measurement they have reflected the spatial and temporal variability of the measurements; when assigned background values, they are constant in space and time. Background values recently recommended by SAI (Reid and Reynolds, 1979) are NO = 1.0 ppb, NO2 = 2.0 ppb, O3 = 40 ppb, HNO2 = 0.1

ppb, H_2O_2 = 0.01 ppb, SO_2 = 10 ppb, CO = 100 ppb, olefins = 0.4 ppb, paraffins = 35 ppb, aldehydes = 5.0 ppb, and aromatics = 6 ppb.

Phenomena such as long-range transport and recirculation of natural emissions may lead to elevated concentrations aloft and at boundaries; thus there is no general upper limit to the concentrations that should be used where there is no measured data even when emissions are low in the boundary region. However, measured concentrations in remote tropospheric areas may provide a lower bound for the boundary concentrations. For populous rural areas (e.g., the northeastern and midwestern U.S.), these lower bounds will likely be too low. Near more isolated areas such as San Francisco or Denver, they may be approached.

Neither the LIRAQ Co's nor the SAI model background values reflect probable remote tropospheric concentrations (e.g., clean marine air or upper troposphere). If remote tropospheric values were to be estimated from the best available data, CO concentrations would be 100-200 ppb (Seiler, 1975), NO concentrations would be 0.004 - 0.020 ppb (McFarland et al., 1978; Drummond 1979), and NO_2 concentrations would be 0.010-0.100 ppb (Drummond, 1979, Noxon, 1978). Remote tropospheric SO_2 concentrations are likely to be below 0.4 ppb (Georgii, 1978), while ozone concentrations of 20-60 ppb seem representative of clean air (Rasmussen, 1976) (ozone and SO_2 concentrations probably increase with altitude because of deposition on surfaces). Background concentrations of hydrocarbons are, as yet, not reliably known. Concentrations of 6-1600 ppb of non-methane and non-aldehydic C_2-C_6 hydrocarbons have been reported (Robinson et al., 1973). Recent estimates of alkenes are of the order of 1-5 ppb, alkane hydrocarbons 5-100 ppb. While no direct estimates of aldehydes are available, model calculations (of the free troposphere) suggest 0.1-1 ppb of CH_2O from methane oxidation, and total aldehydes of 0.5-5 ppb. H_2O_2 seems likely to be present at 0.1 to 1 ppb, while HONO is unlikely to be above 0.03 ppb (Wuebbles et al., unpublished work, 1979).

One method of estimating boundary concentrations might be to run "nested models"; i.e., to first run a larger scale model and use the predictions to generate boundary conditions for a smaller scale model. While conceptually attractive this concept has received only rather limited testing thus far.

Future Year Simulations

If boundary conditions reflect natural background concentrations or low level emissions from areas unlikely to experience significant growth or emissions reductions, there is no reason to adjust boundary conditions in future year applications, and indeed, boundary conditions have usually not been adjusted in model applications. (In the case of LIRAQ applications the portion of the boundary concentration that reflects the concentration just within the boundary is automatically adjusted by variations in that concentration as they involve emissions changes).

However, if emissions from another area, or from developing rural areas, play a significant role in determining either baseline or future boundary concentrations, then a satisfactory treatment of boundary conditions should reflect projected emissions in nearby areas (nearby might be as far away as several hundred km for some pollutants, e.g., CO, 03 and SO2). Unfortunately because of the large amount of effort involved, a detailed treatment will not likely be attempted except under extraordinary circumstances.

Summary

In summary, initial and boundary concentrations can have an influence on model-computed results for suburban and downwind areas that approaches the importance of emissions, and can have a significant influence on calculations of secondary pollutants even in central urban areas. Moreover, this is true even when the initial and boundary concentrations are of the order of a few to one hundred ppb. On the other hand boundary and initial concentrations will rarely if ever cause or make a major contribution to violations of standards for primary pollutants. The problem with respect to concentrations approaching the standards is largely limited to secondary pollutants.

In the absence of a special field data collection study, there are almost no data available that are really useful in guiding the choice of boundary conditions or of conditions aloft. Data obtained with normal air quality monitoring instruments are so imprecise that they are often difficult to interpret.

Although most current model applications use lower background values than were used even two or three years ago, currently used concentrations are not low enough to be a negligible factor in model calculations. Moreover, except for carbon monoxide, there is little reason to believe that even the lower concentrations used in model applications are as low as those representative of a remote tropospheric air mass. Conversely, there is also little reason to believe that the air 10-100 km from a representative urban center should be very close to remote tropospheric air in its composition. Thus, the specification of initial and boundary conditions remains a major problem in the verification and application of air quality models, and all proposed solutions to the problem are of an essentially *ad hoc* nature.

The problem of choosing initial and boundary conditions should be considered when the choices of model domain and meteorological conditions are made. If a city is at the western edge of a populous region, it is likely that boundary concentrations will be lower and more easily estimated so that calculations will be more reliable if the winds are from the west. If a city is embedded in a megalopolis, a large-scale model is likely to be required at least for species like ozone that require several hours for formation from their precursors.

SPECIAL FIELD MEASUREMENT STUDIES

The purpose of a special field study is to augment the existing emissions and aerometric data gathering efforts toward the end of providing an adequate data base for preparing model inputs, evaluating performance, and examining alternative control strategies. Because provision of model input and evaluation data is not the normal objective of routine emissions and aerometric monitoring, one invariably encounters a data base that is sparse in one or many components when applying a model to an urban area. As might be expected, there is wide variation in the number and type of measurements made in urban areas in the United States. All cities have at least a few surface wind stations, for example, but the amount of data on upper level winds and atmospheric stability varies greatly. The number of air quality monitors likewise varies and, to a degree, reflects the predominant air quality concern in each city. Oxidant monitoring in St. Louis and Los Angeles is extensive, whereas in Las Vegas, for example, concern seems to focus on carbon monoxide. Routine measurements are not made of concentrations of hydrocarbons by species or pollutant concentrations aloft, though these measurements are made during some special field studies. Finally, emissions inventories for various cities range from region-wide estimates of nominal emissions rates for various pollutants (in the NEDS format) to detailed, gridded inventories that disaggregate emissions according to individual sources or source categories.

A field observation program can be expensive and care must be taken in its planning and coordination to ensure that the desired information is obtained. The costs and benefits of long-range vs short range scheduling are an important consideration. For example, suppose one assigns a very high priority to obtaining supplemental data on the worst air pollution day of the year. This could be guaranteed by taking measurements every day during the smog season - but at very high cost. Or, one might schedule field programs long in advance for several days during the peak smog season in the hope that the worst day happens to fall on one of those days. A third alternative might be designed to be implemented on very short notice - perhaps no more than a day or two. Short notice would be necessary because adverse conditions cannot be forecast reliably more than a day or two in advance. This approach, while having a fairly high probability of success, also has inherent drawbacks. First, because the worst day cannot be known in advance, the field program will have to be implemented on several of the most adverse days. Second, the requirement for deployment of equipment and personnel on short notice will preclude the use of some observation systems that require long lead times for set-up. Personnel costs per unit time will also tend to be higher.

The following sections discuss some of the possible components of a special field measurement study; previous studies carried out in the San Francisco Bay Area and Los Angeles are also summarized.

Meteorological Data

Supplementary meteorological data gathering efforts are commonly required to upgrade the data base available from routine monitoring activities. As previously indicated, measurements of upper air conditions frequently are sparse and do not provide enough information for satisfactory interpolation or for use in windfield or mixing depth models. Moreover, in situations where special conditions prevail, such as topographical influences on air flows, routine measurement networks of surface conditions can prove to be inadequate to supply necessary detail in the input data.

Wind data (speed and direction) are normally available from surface measurement networks. However, additional observations may be needed to adequately characterize wind flows in outlying or suburban areas or in regions where significant terrain features may substantially alter flow patterns. The principal supplementary activity, as pointed out above, will be the collection of vertical wind soundings to obtain upper air wind data. Such measurements can be effected by releasing and tracking pilot balloons, or pibals, at various locations and times of the day. An alternative method that can be used for measurements up to a few hundred feet is to mount instruments on fixed towers.

Radiosondes are instrumented balloons that are released and tracked to give information on the temperature structure of the atmosphere, pressure gradients, relative humidity as a function of elevation, and wind speeds and direction. However, because they are considerably more expensive to deploy than pibals, their use is primarily in connection with the first three variables mentioned and only incidentally for wind measurements. Aircraft are also used to measure vertical temperature structure and pressures; these data are used to infer mixing depths throughout a region. An aircraft has the advantage that it can be flown according to a fixed pattern, such as a vertical spiral, and thus make a measurement at a predetermined location. Another method for measuring the depth of the inversion layer is by means of acoustic soundings, which are made from ground-based stations. Finally, tetroon releases, when tracked from the ground or from aircraft, can be used to follow air parcel trajectories. They have the advantage over pibals that, because they follow isentropic surfaces, they can be tracked for longer distances.

Air Quality Data

Most locations will have an existing network for collecting air quality data; a routinely collected data base, however, may fail to satisfy the needs of a particular study. The data may be deficient in terms of their spatial or temporal coverage, lack of upper air measurements, or lack of detailed information on some pollutants of interest.

Should the existing network be judged too sparse, it may be supplemented by additional fixed stations or by mobile stations. Mobile stations are probably more cost-effective in a supplementary data gathering effort because they are much more adaptable and can be converted to other tasks after the study has been completed. Sample hydrocarbon analyses can be carried out at a central location if bag samples are collected and transported to the analysis facility.

Routine networks will often fall short of model requirements with respect to the collection of boundary condition data. These data must be collected in outlying areas and aloft. Boundary condition requirements vary according to the region being modeled and the conditions on the day being modeled. If large pollutant fluxes are present on the upwind boundary, boundary conditions will be important; if air is coming from an area of minimal concentrations (e.g., the Pacific Ocean in the case of San Francisco), boundary conditions may not be needed in much detail. Initial conditions also are important and may require supplementary data collection. Accurate specification of initial conditions can be the key to a satisfactory simulation.

Routine monitoring networks frequently do not provide concentrations of particular pollutants of interest. The analyses that are almost invariably lacking are the concentrations of individual hydrocarbon species, knowledge of which is necessary for a photochemical model. The different hydrocarbon species undergo different reactions in the kinetic mechanism, and the course of the overall chemical transformation is highly dependent on the hydrocarbon distribution. Thus, it frequently is necessary to upgrade the analytical capabilities of existing monitoring stations during special field studies.

In summary, it is almost invariably necessary to supplement routine data collection by conducting a special study. The actual extent of the additional activities must be determined after comparing the existing data base with the data needs of the model. Costs of this supplementary data collection are difficult to estimate in general. Miedema et al., (1973) surveyed monitoring costs in the early 1970s, and their report gives capital and operating costs for many instruments and types of measurements. However, it must be recognized that their estimates have been overtaken by inflationary cost increases; present-day costs will be substantially higher. An updated monitoring cost survey is in preparation (Lutz, 1979), and a report may be forthcoming in 1979 giving more current figures.

Even in the most widely monitored areas the routinely collected data are usually insufficient to characterize the basic model inputs with the desired spatial and temporal resolutions. Determining the spatial variation of the mixing depths and pollutant concentrations aloft will almost always require a special field study. The quantity of data collected and the duration of the field study should depend on both the available resources and the complexity of the modeling region. The following sections describe two field studies that were undertaken to supply data for airshed modeling.

Collecting Supplementary Field Data for LIRAQ

Data bases of meteorological information and pollutant concentrations adequate for applying the LIRAQ model to San Francisco were not available, even though the BAAQMD had, for a number of years, operated one of the most extensive networks of meteorological and air quality monitoring stations in the nation. Even with the addition of data from other existing sources in the Bay Area, it appeared that the resulting data base would be inadequate. Accordingly, a field data collection and processing effort was carried out. In addition to data from the existing surface measuring stations of the BAAQMD and other agencies, data were collected by temporary surface stations deployed during selected field intervals and by an aircraft operated by the NASA Ames Research Center (ARC) over predetermined flight paths. The field program also included a fluorescent tracer particle study and a correlation spectrometer study, each designed to test some of the assumptions on which LIRAQ was based.

Field Observations at Temporary Sites. To augment a single routine meteorological rawinsonde site at Oakland International Airport, a group of sites was selected for radiosonde and pibal observations of vertical profiles of wind, temperature, and humidity, using portable equipment. The California State University at San Jose, under subcontract to the NASA-Ames Research Center, was responsible for setting up and operating these portable sites. The University agreed to provide four teams for the primary purpose of taking vertical soundings and making wind observations. Generally, one team operated the rawinsonde equipment at the University, and other teams performed mobile activities in the field. One of the teams operated a mobile radiosonde van on loan from NOAA/National Weather Service and doubled as a pibal team. The remaining teams generally took observations and all of the teams took surface meteorological observations. In addition to these teams, three portable wind observation stations for specialized observations and the services of one or two mobile vans were available from the BAAQMD.

Observations by Instrumented Aircraft. The bulk of vertically resolved air quality and meteorological data needed for model development was provided by an instrumented twin-engine Cessna aircraft flown by the ARC. Vertical soundings were taken by the aircraft within the polluted surface layer at specified sites and along climbing and descending flight paths between the sounding sites. Parameters sampled included wind speed and direction, temperature, dew point, ozone, total oxidant, oxides of nitrogen, carbon monoxide, sulfur dioxide, and hydrocarbons. Sampling was instantaneous at 20 second intervals.

Barringer COSPEC Studies. During the first year of the project, in conjunction with the other field observations, a program of special studies involving the use of the Barringer correlation spectrometer (COSPEC) was carried out. The program involved obtaining measurements of the integrated burden of nitrogen dioxide over a straight-line path in combination with fixed-point surface measurements of total oxides of nitrogen dioxide and total sulfur.

The purpose of these studies was to provide spatially-averaged concentration data for use in testing the spatially-averaged assumptions embedded in the LIRAQ model.

Flourescent Tracer Studies. In the second project year, a set of specialized fluorescent tracer experiments was designed and carried out to examine the questions of transport through the interface between the marine air and the overlying inversion layer. That information was necessary to assess the assumptions made in the submodel for determining a mass-consistent wind field.

In conjunction with each of the experiments (which were not conducted on normal field experiment days), temperature soundings were taken on a regular basis by volunteer aircraft of the Oceanic Society in order to define the structure of the inversion layer.

No tracer was found to penetrate the lower interface of the inversion layer during either of the two experiments. These results neither confirm nor reject the hypothesis of transport through the inversion interface, but they do reinforce the widely held assumption that such penetration is unlikely, especially when the inversion height is not changing. Meteorological and time constraints unfortunately prevented experiments under inversion destruction conditions such as breaking waves and thermal ablation of the inversion layer. The original design concept had incorporated experiments of that type.

CALTRANS Field Study in Los Angeles

The California Department of Transportation (CALTRANS) has evaluated the SAI model using data collected on 26 June 1974. The data available for this day were determined to be insufficient to fully describe the three-dimensional wind field and the distribution of pollutants aloft, so a study was undertaken to collect additional data during the 1975 smog season to use in characterizing the aloft wind and pollutant patterns. This study was a multiagency cooperative effort involving CALTRANS, the CARB, air pollution control districts in five counties, NOAA, the EPA-Research Triangle Park, the National Environmental Research Center-Las Vegas, the U.S. Navy, and the University of California at Los Angeles.

The sampling program consisted of the following elements:

- Instrumented aircraft flights.
- Pibal launches (for winds aloft) throughout the basin.
- A network of mechanical weather stations sited to obtain an adequate representation of the surface wind field for the region.
- Solar radiation measurements both above and below the inversion layer and covering the modeling region.
- Air quality data from all the measuring sites in the regions.

To reduce expenses, sampling was performed only when the San Bernardino County APCD predicted an episode that would exceed 690 ug/m3 (35 pphm) oxidant. Although the APCD predictions were not always accurate, this procedure did result in some very satisfactory data sets. The total cost of obtaining these data, including editing and entering them into a computer, was about $150,000. No model evaluation runs have yet been done using these data.

REFERENCES

Andersen, G.E., S. R. Hayes, M. J. Hillyer, J. P. Killus, and P. V. Mundkur, "Air Quality in the Denver Metropolitan Region 1974-2000." Preared for the U.S. Environmental Protection Agency, Region VIII by Systems Applications, Incorporated, under contract number 68-01-4341, May 1977.

Booz-Allen and Hamilton, "Regional Emission Projection System," prepared for U. S. Environmental Protection Agency, EPA-450/3-74-051, 1974.

Dickerson, M. H. (1978), "MASCON - A Mass Consistent Atmospheric Flux Model for Regions with Complex Topography," *J. Appl. Meteor.*, Vol. 17, No. 3, pp. 241-253.

Drummond, J. W., private communication, 1979.

Duewer, W. H., M. C. MacCracken, and J. J. Walton, "The Livermore Regional Air Quality Model: II. Verification and Sample Application in the San Francisco Bay Area," *J. Appl. Meteor.*, *17*, 273-311, 1978.

Georgii, H. W., "Large Scale Spatial and Temporal Distribution of Sulphur Compounds," *Atmos. Environ.*, *12*, 681-690, 1978.

Leong, E. Y. and R. Y. Wada, "Emission Inventory Projections: Hindsight, Insight and Foresight," proceedings of the Air Pollution Control Association Specialty Conference on Emission Factors and Inventories, Anaheim, California, November 13-16, 1978.

Lutz, D. (1979), Environmental Protection Agency--personal communication to Systems Applications, Incorporated.

MacCracken, M. C. (1975), "User's Guide to the LIRAQ Model: An Air Pollution Model for the San Francisco Bay Area," UCRL-51983, Lawrence Livermore Laboratory, Livermore, California.

MacCracken, M. C., and G. D. Sauter, editors, "Development of an Air Pollution Model for the San Francisco Bay Area, Volume I." Final report prepared for the U. S. Energy Research and Development Administration by Lawrence Livermore Laboratory under contract number W-7405-Eng-48, October 1, 1975.

McFarland, K, D. Kley, W. C. Kuster, A. L. Schmeltekopf and J. W. Drummond, "NO, 03, JNO2, and CO Measurements Made in the Equatorial Pacific Region," paper given at the 1978 Fall American Geophysical Union Meeting.

Miedema, A. K., et al. (1973), "Cost of Monitoring Air Quality in the United States," EPA-450/3-74-029, Environmental Protection Agency, Research Triangle Park, North Carolina.

Noxon, J. F., "Tropospheric NO2," *J. Geophys. Res.*, *83*, 3051-3057, 1978.

Perardi, T. E., et al., "Preparation and Use of Spatially and Temporally Resolved Emission Inventories in the San Francisco Bay Region, Journal of the Air Pollution Control Association, Vol. 29, No. 4, pp. 358-364, April 1979.

Rasmussen, R. A., "Surface Ozone Observations in Rural and Remote Areas," J. Occupational Medicine, 18, 346-350, 1976.

Reid, L. E. and S. D. Reynolds, "The Conversion of LIRAQ Inputs to SAI Airshed Model Inputs," SAI Report EF79-50, April 1979 (Draft).

Robinson, E., R. A. Rasumussen, H. H. Westberg, and M. A. Holdren, "Non-urban Non-methane Low Molecular Weight Hydrocarbon Concentrations Related to Air Mass Identification," J. Geophys. Res., 78, 5345-5351, 1973.

Seiler, W., "The Cycle of Atmospheric CO," Tellus, 26, 116-135, 1974.

Tesche, T. W. (1978), "Evaluating Simple Oxidant Prediction Methods Using Complex Photochemical Models (Monthly Technical Progress Narrative No. 1)," EM78-14, Systems Applications, Incorporated, San Rafael, California.

U. S. Department of Transportation, Federal Highway Administration "Special Area Analysis - Part 4. Special Area Pollution," prepared by L. R. Seiders, Comsis Corporation for FHWA Urban Planning Division, Washington, D. C., August 1973.

U. S. Department of Transportation, Federal Highway Administration, "SAPOLLUT/SAPLSM User's Guide," FHWA Urban Planning Division, HHP-23, Washington, D.C., October 1976.

U. S. Environmental Protection Agency, "Guidelines for Air Quality Maintenance Planning and Analysis, Volume 7: Projecting County Emissions," EPA-450/4-74-008, Research Triangle Park, North Carolina, May 1975.

U. S. Environmental Protection Agency, "Mobile Source Emission Factors," Office of Transportation and Land Use Planning, Washington, D. C., EPA-400/9-78-005, March 1978.

Wuebbles, D. J., W. H. Duewer, R. Tarp, and J. S. Chang, unpublished recent calculations, 1979.

4. THE EVALUATION OF PHOTOCHEMICAL MODEL PERFORMANCE

The purpose of this section is to consider those efforts involved in the assessment of photochemical air quality model performance. In Section 2, a number of different photochemical models were identified that might potentially be employed in the preparation or revision of State Implementation Plans (SIPs). Given this variety of models, there is a need for an adequate understanding of their performance, both in a relative sense, compared with other models, and in an absolute sense, for a particular application. This understanding will enable the user to choose an appropriate model to ensure that it performs adequately in the intended application.

In the past, model performance was usually evaluated by the model developer in the course of an application or a special evaluation study. Although air quality model usage has been increasing over the last several years, there are no EPA-published guidelines describing how such studies should be carried out. Recognizing this deficiency, EPA has recently commissioned a study to examine performance measures and standards as well as general procedures for conducting a performance evaluation study. One of the resulting reports, by Hayes (1979), considers alternative means of measuring model performance and possible ways in which performance standards might be established. A companion report by Hillyer, Reynolds, and Roth (1979) presents a generalized, step-by-step procedure that may be used to evaluate the performance of an air quality simulation model. The discussion in this section adopts many of the concepts presented by Hayes (1979) and Hillyer, Reynolds, and Roth (1979). The reader is referred to these reports for background and supplemental information.

Some of the terminology used to describe the model evaluation process may require clarification. The phrases "model validation" and "model verification" are frequently used to designate the process of comparing model predictions with suitable observations. In this section, however, the terms "validation" and "verification" are avoided in referring to the evaluation of a model. Instead, the more general phrase "model performance evaluation" is used since it is more representative of the process that this section describes. The "validity" of a model is taken to be a concept defining how well model predictions would agree with the appropriate observations, given a perfect specification of model inputs. That is, validity relates to the inherent quality of the model formulation. The term "verification" is reserved to describe a successful (or positive) outcome of the model evaluation process.

The examination of model performance is motivated by two factors: First, the model treatments of various physical and chemical phenomena usually involve a number of approximations; second, the information provided as inputs is often subject to considerable uncertainty. The principal sources of these uncertainties include (Seinfeld, 1977):

- Wind velocity components.
  - Uncertainties in available wind speed and direction measurements.
  - Inadequate number of nonrepresentative locations of measurement sites (especially aloft).
  - Approximations associated with wind field analysis techniques used to prepare model inputs.
- Source emissions and removal functions.
  - Inaccurate specification of source location.
  - Inaccurate estimation of plume rise.
  - Errors in emission factors for stationary sources.
  - Inadequate representation of actual driving characteristics in the federal mobile emissions test procedures.
  - Errors in emission factors and vehicle miles traveled for mobile sources.
  - Inaccurate characterization of temporal variations.
  - Inadequate parameterization of pollutant removal mechanisms.
- Chemical reaction mechanism.
  - Omission or inadequate characterization of chemical reaction steps.
  - Uncertainties in measurement or specification of reaction rate constants.
  - Inaccurate characterization of temperature effects.
  - Uncertainties associated with categorizing species into reactive groups (such as paraffins, olefins, aromatics, etc.).
- Initial and boundary conditions.
  - Inadequate spatial characterization of the concentration field on the upwind boundary of the region.
  - Inadequate characterization of concentrations aloft.
- Numerical solution methodology.
  - Computational errors associated with the use of finite difference methods.
  - Computational errors associated with other numerical techniques.

In addition, it should be noted that the air quality data employed to judge model performance are subject to error. Instrumental errors can occur, or the spatial character of the measurement may not be directly comparable to that of the model predictions (e.g., a point measurement may be compared to a spatially averaged model prediction). Thus, there are a number of sources of error in even the most sophisticated photochemical models. Some of the simpler techniques, such as trajectory models, can be derived from the same governing equation as grid models with appropriate assumptions, and so they are subject to as many, if not more, sources of error than are the more sophisticated techniques. The purpose of the model evaluation study is to test the entire model, including formulation, available aerometric and emissions data, and model input preparation procedures, to obtain some quantitative measure of model performance. Hilst (1978) suggests in a recent report the additional need to evaluate the performance of specific components of a model. For example, special studies could be performed to assess the plume rise algorithm included in a Gaussian point source model. These types of investigations will be especially important in the initial applications of a model because they may aid in the identification and rectification of inadequate treatments of atmospheric processes.

OUTLINING THE MODEL EVALUATION STUDY

The initial outline for a model performance evaluation can be divided into two parts:

- Specific definition of the scope of the study;
- Selection and design of performance measures and standards (i.e., performance evaluation criteria);

Defining the extent of required verification entails addressing the following issues: size and boundaries of the modeling region, model resolution, pollutants to be included, time period to be simulated, and type and number of meteorological regimes.

Definition of the Size and Boundaries of the Modeling Region

The appropriate modeling region, a function of the model and of its application, should contain all areas with significant population and those where the peak concentrations are expected to occur. Trajectory models are additionally constrained by the distances over which the basic model concept is valid. Grid models are constrained by the number of grid cells that can be accommodated in the available computer core storage. These cells must be judiciously arranged to cover the area of interest. If the entire area cannot be accommodated by the host computer, the grid cell size can be adjusted, though this adjustment of course influences the resolution of the model.

Definition of Spatial and Temporal Model Resolution

Spatial resolution (i.e., spacing between grid points) is related to region size and the number of grid cells into which the region is divided. In addition to the amount of computer storage, the number of grid cells is constrained by time limitations. For example, though the SAI Airshed Model has no inherent limitations on the number of grid cells, doubling the number of cells slightly more than doubles the necessary computer time.

In addition, the spatial resolution of the model is influenced by the spatial resolution of the input data. Assuming the existence of a uniformly spaced wind measurement network and invoking sampling theory concepts, Lamb and Seinfeld (1973) argue that, theoretically, a model cannot resolve features in the initial concentration field on a scale smaller than one-half the distance between the monitoring stations. Furthermore, the model cannot resolve features in the predicted concentration field on a scale smaller than the resolution of the emission inputs. In previous photochemical model applications, the grid specification has usually been more heavily influenced by the resolution in the available emissions data than by the spacing between the wind stations. This situation does not appear unreasonable when one considers that re-creation of the actual flow field is not as important in the evaluation of future emissions patterns as it is in the attempted simulation of an historical air pollution episode. In general, the grid spacing should be sufficiently fine to resolve the expected characteristics of the concentration field as well as the important features of the physical and chemical atmospheric processes. The horizontal grid spacing employed in previous grid modeling studies has ranged from 1 to 10 kilometers.

Grid models are best suited for predicting the concentrations of species, such as ozone, that have a one-hour air quality standard. Averaging times of up to 24 hours can be accommodated with grid models, but longer time intervals would require too much computer time to be practical. If the one-hour averages calculated by the model are to reflect actual conditions during that hour, the input data should also have a comparable temporal resolution. This requirement is discussed in more detail for each input variable in Section 3.

Time Period to be Simulated

The simulation must include the induction period for the formation of photochemical pollutants. To minimize the effects of the initial conditions, the simulations should be started prior to the morning rush hour. In some cases pollutant carryover from the previous day is important, making a multiple day simulation desirable. The length of the simulation depends on the time period over which significant pollutant concentration levels are observed or predicted, e.g., from early morning to late afternoon.

Choice of Meteorological Conditions

In general, performance is evaluated by applying the model to one or more historical air pollution episodes. If adequate performance is achieved, the model can then be assumed ready for use in calculating the effects on air quality levels of changing emissions. Ideally, the model's ability to predict the effects of emissions changes should be directly assessed. However, historical data bases suitable for evaluating model performance usually cover only a limited span of time over which emissions have not significantly changed. A limited check on this aspect of model performance can be obtained by carrying out simulations for both weekdays and weekends. (However, the value of this check is limited by the generally low quality of weekend emissions inventories.) In addition, as models are applied to several urban areas having varying emissions intensities, some information will be obtained on the model's prediction capabilities under different emissions conditions. Until control strategies are actually implemented and their effects on air quality are measured, model performance must necessarily be evaluated in the context of present emissions conditions.

An air pollution episode suitable for use in a control strategy evaluation should possess three main attributes. First, it should not place undue "stress" on the model such as would be the case for episodes characterized by extremely complicated wind flows or unusual ground-level pollutant concentration distributions. The objective is to select an episode for which the model can be expected to perform well; use of an unduly complicated episode increases the likelihood of degraded model performance and attendant uncertainties about the meaning of the strategy evaluation simulations. Second, the episode should stress the control strategies: if the aim is to evaluate the effects on the exposure and dosages experienced by the human population of controlling point source emissions in a city with many point sources in its eastern portion, an episode with predominantly westerly winds should not be chosen. Third, the episode should represent worst-case conditions: the meteorological and emissions conditions should be conducive to the occurrence of upper percentile pollutant concentrations.

Selection of an episode is a screening process. First, historical data are searched to define periods of from one day to several consecutive days during which elevated pollutant concentrations were observed, preferably at several monitoring locations. Then, the extent of data available, from both the routine monitoring network and special studies, is determined. Those episodes for which serious data gaps occur are deleted from further consideration. Finally, the remaining candidate episodes are examined individually to select the most promising episode(s).

Determination of the Number of Meteorological Regimes

Often it may be advantageous to consider the selection of more than one episode. If high concentrations are observed under very different meteorological regimes, it may be necessary to test alternative emissions control strategies under the various worst-case regimes. Of course, the number of episodes that can be analyzed is limited by the resources available to prepare data bases and to carry out the required model evaluation exercises. Two basic considerations in selecting the number of regimes to be included in an evaluation study are:

- Specification of a sufficient number to enable characterization of model performance.

- Selection of regimes that will be appropriate for use in the applications studies.

In general, two to six different meteorological scenarios are recommended. A final control strategy should be assessed using at least two or three different "worst-case" meteorological conditions to determine compliance with the NAAQS. We refer the reader to Section 6 for a further discussion of the choice and usage of "worst-case" conditions.

Performance Measures and Standards

No measured quantity is equivalent to a photochemical model's concentration predictions; therefore, no absolute measure of model performance exists. The best available comparison is between station measurements and grid averaged predictions. That is, it is necessary to assume that station measurements are representative of the average concentration throughout the grid. Although there is no officially accepted method of comparing observed and estimated concentrations, several techniques are commonly used, including calculating correlation coefficients, percent differences, and standard deviations.

Hayes (1979) has carried out one of the most systematic considerations of how to measure model performance to date. Table 4-1 lists the measures and standards from this study that are thought to be most useful for photochemical models. Hayes (1979) has identified three objective rationales for setting model performance standards, viz:

- Health effects
- Control level uncertainty
- Guaranteed compliance.

The first of these contains the premise that computed considerations should not deviate from the true values sufficiently to result in significantly underestimated health effects. In the second rationale, uncertainties in the percentage of emissions control required should be held to tolerable values. The third rationale implies that compliance with air quality regulations should be "guaranteed," that is, that uncertainty in the model's predictions should be biased in a conservative direction.

Unfortunately, these rationales, which state desirable goals for model evaluation, cannot be applied at the present time because there is no objective basis for setting the required uncertainty levels. It is therefore recommended that a "pragmatic/historic" rationale (Hayes, 1979) be employed to evaluate the performance of photochemical models. In this method, the performance standard is that the model must perform at least as well as in recent evaluation efforts carried out for a similar urban area and model application.

Trajectory models present a special case in determining performance measures and standards. Because trajectory models yield predictions along a trajectory path, special attention must be given to how the computed and measured values are to be compared. Two techniques used in previous model performance evaluation studies include:

- Running the trajectory near stations of interest and estimating the "measured" concentration at the trajectory location through interpolation of the available observations (Eschenroeder, Martinez, and Nordsieck, 1972).

- Defining a set of trajectories that will arrive at a monitoring station location throughout the period of interest.

Table 4.1. Sample Photochemical Model Performance Measures and Standards

Performance Attribute	Performance Measure	Performance Standard
Accuracy of the peak prediction	Ratio of the predicted station peak to the measured station peak (could be different stations)	Limitation on uncertainty in aggregate health impact and pollution abatement costs*
	Difference in timing of occurrence of station peak†	Model must reproduce reasonably well the phasing of the peak--say +1 hour
Absence of systematic bias	Average value and standard deviation of the mean deviation about the perfect correlation line, normalized by the average of the predicted and observed concentrations, calculated for all stations during those hours when either the predicted or the observed values exceed the NAAQS	No or very little systematic bias at concentrations (predictions or observations) at or above the NAAQS; the bias should not be worse than the maximum bias resulting from EPA-allowable calibration error (-8 percent is a representative value)
Lack of gross error	Average value and standard deviation of the absolute mean deviation about the perfect correlation line, normalized by the average of the predicted and observed concentrations, calculated for all stations during those hours when either the predicted or the observed values exceed the NAAQS	For concentrations at or above the NAAQS, the error (as measured by the overall values of the average and standard deviation of the mean normalized absolute deviation) should not be worse than the error resulting from monitoring instrumentation error
Temporal correlation†	Temporal correlation coefficients at each monitoring station for the entire modeling period and an overall coefficient averaged for all stations	At a 95 percent confidence level, the temporal profile of predicted and observed concentrations should appear to be in phase (in the absence of better information, a confidence interval may be converted into a minimum allowable correlation coefficient by using an appropriate t-statistic)
Spatial alignment	Spatial correlation coefficients calculated for each modeling hour considering all monitoring stations, as well as an overall coefficient average for the entire day.	At a 95 percent confidence level, the spatial distribution of predicted and observed concentrations should appear to be correlated

---

* These measures are appropriate when the chosen model is used to consider questions involving pollutants subject to short-term standards. They are most important when the pollutant is also received.

† These may not be appropriate for all regulated pollutants in all applications. When they are not, standards derived based on pragmatic/historic experience should be employed.

Source: Hayes (1979).

> When the air parcel is within a certain distance of the station location (say 1 to 5 kilometers), then the predictions can be time averaged and compared with the observations. In some instances, it may be necessary to combine the results from several trajectory runs to calculate a time-averaged prediction commensurate with the measured values (Wayne, Kokin, and Weisburd, 1973).

These two techniques may be combined in evaluating the performance of a trajectory model. In general, the second scheme cited above is preferable for establishing the model's performance characteristics since it is not subject to the added uncertainties introduced by the use of the interpolation procedure.

In summary, regarding both grid and trajectory models, there is a clear need for further studies to establish suitable model performance measures and standards.

MODEL EVALUATION PHASE

The fourth and final phase of the model evaluation effort entails the adaptation of the model to the study area, the collection of supplemental data, the preparation of inputs, the exercise of the model, the analysis of the results, and the rectification of performance deficiencies. The following sections discuss the most important topics related to these issues.

As part of the model selection task, the issue of the availability of a suitable computer was raised. At this point, it should be emphasized that achieving successful runs of the programs on a computer for which they have not been previously implemented can be very time consuming, possibly requiring assistance from the model developer. To ensure that the programs are transferred correctly, a test case should be supplied by the model developer or past users to be employed to test the programs. The programs should be up and running on the user's computer before any modifications are made.

Adaptation of the Model to the Study Area

Model modifications fall under two categories: modifications to accommodate the modeling region and alterations in the model algorithms. The first of these types of modifications may simply require routine alterations to the codes. For example, in some models the array dimensions must be changed for each modeling region of a different size or shape. This is usually a fairly simple and straightforward process.

The second type of modification can be somewhat more complex. The best available model may not treat all the atmospheric processes as the user would like. In addition, the available procedures for preparing model inputs may not be the most suitable for use in the new study area. Thus, it may be necessary to modify an existing algorithm in the model or, perhaps include a process not treated by the model. For example, the particular plume rise algorithm contained by the model may not be considered to be the best one for the intended application. A better-suited algorithm can be coded and inserted in the model as a replacement for the existing algorithm. To make this kind of modification requires a thorough knowledge of the computer code and may require the assistance of the model developer. After the modifications are made, the program should be carefully tested to determine whether the new algorithm is working properly and to assure that no other operations have been changed accidentally.

Collection of Model Inputs

The methods of obtaining the data for specifying the model inputs are described in detail in Section 3. As the data are gathered, they must be checked for consistency and reformatted for use with preprocessor programs, which then use the collected data to construct the input files for the main simulation program.

The conversion of the raw data into the simulation program input files is one of the most critical tasks in the evaluation study. The preprocessor programs should make the maximum use of the data, and the output from these programs should be carefully reviewed to ensure that the results meet the user's expectations. Since the data preparation process is very dependent not only on the quantity of data gathered, but also on the procedures employed to estimate the inputs, a good understanding of the algorithm and assumptions used by the data preparation programs is essential to prepare a set of adequate model inputs.

Performance of Model Simulations

Once the input files have been prepared it is a simple task to exercise the simulation program. The model results should be saved on a permanent file (e.g., a tape) so that performance measures can be estimated and subsequent computations can be carried out by the user. If the run is successful, other analyses may be of value, such as dosage based on the distribution of pollutant concentration levels and the population, concentration isopleths, areal extent of the region for which concentrations exceed the NAAQS, and concentration frequency distributions.

Assessment of Model Simulation Results

The first step in analyzing the results of the model simulation is to compare each computed performance standard. If the measure meets the standard, the user may proceed with some confidence that the model will produce reliable results for the intended applications.

If a performance measure fails to meet a performance standard in some respect, further analysis is indicated, as outlined below. It should be noted that even if the measure does meet the performance standard, this type of analysis will give much useful information about model behavior; therefore, the analysis is strongly recommended in all cases.

The analysis of the model evaluation results should center on the differences between computed and observed pollutant concentrations, that is, the residuals. These residuals can arise from three sources:

- Errors in the input data (emissions data, meteorological data, initial and boundary conditions).
- Errors in the formulation of the model (approximations made in modeling pollutant transport or chemical transformations).
- Errors in the air quality measurements used for comparison.

Estimates of the precision of various aerometric measurements can be determined by replication of the measurements and by the application of standard statistical techniques. The accuracy of the measurements can be characterized through an analysis of instrument calibration data. Nonzero residuals can also result from the extrapolations and interpolations necessary to generate a complete model input data set from insufficient data. The errors introduced by these extrapolations and interpolations are somewhat more difficult to quantify than are instrumental errors. In addition, station exposure to localized biases such as near roadway sources of NO may unduly influence the measurements.

Discrepancies introduced by shortcomings in the model's formulation are difficult to evaluate because there is no "true" model of the relevant atmospheric processes for comparison. Also, in light of the uncertainties of the input and comparison data, error due to model formulation cannot be isolated from the total modeling process.

Careful analysis of the residuals can yield much useful information about the model, even if quantitative statements about sources of error cannot be made. Several different ways of analyzing residuals can be informative [see, for example, Koch and Thayer (1971)]. Plots of residuals against time of day can reveal systematic biases, which might result from an inadequate kinetic mechanism in a photochemical model. Dependence of the magnitude of residuals on concentration might indicate that a monitor is poorly located to detect a large area-wide concentration level, or that the wind field inputs for the model incorrectly represent the transport of a plume from a large point source. Differences between residual dependencies for primary and secondary pollutants can be used to infer deficiencies in the kinetic mechanism or dispersion processes. The examples given here are only a sampling of possible situations; this type of analysis should be guided by the particular situation and a knowledge of the various technical features of the model. We refer the reader to the report by Liu et al. (1976), which includes an evaluation of the performance of three photochemical models based on an analysis of residuals.

Other statistically based analysis methods that can be used to study the evaluation results are referenced below:

- Scatter plots of observed and computed considerations (Anderson et al., 1977, Duewer et al., 1978, MacCracken et al., 1975)

- Correlation between observed and computed results (Reynolds et al., 1979, Duewer et al., 1978, MacCracken et al., 1975)

- Nonparametric tests of location to indicate possible bias of computed concentrations relative to observations (Lehmann, 1975).

Rectification of Inadequate Performance

When the model performance evaluation indicates less than satisfactory performance, diagnostic studies should be carried out to obtain evidence of deficiencies in the model's formulation or its inputs. The model user may need to perform these diagnostic analyses at three points:

- During the initial model adaptation to the study region and preparation of inputs.

- Just after the first unsuccessful model exercise.

- Immediately prior to carrying out the final simulation.

In general, the emphasis of the requisite analyses for each of the three stages is somewhat different.

When the model is initially being adapted to the study region, the user is dealing directly with issues of the adequacy of the data base for use in preparing model inputs. After the user has selected a proposed input preparation algorithm and has exercised the appropriate computer program(s), the resultant input file must be examined to determine whether the atmospheric phenomena represented on the file are adequately characterized. For example, estimating the surface wind field using an inverse distance weighted interpolation scheme can sometimes yield a "jump" in the wind velocity at those locations near the radius of influence of a monitoring station. To correct for this unrealistic wind behavior, it may be appropriate to change the radius of influence input to the program, to employ a smoothing algorithm, or possibly to select an alternative wind field preparation technique. Each file prepared for input to the model must be scrutinized carefully. Previous experience in using the available preprocessor routines is often useful in the anticipation of problems that are likely to occur in using particular algorithms in conjunction with sparse observational data. This experience will be helpful in reducing the number of iterations required to achieve a reasonable initial set of model input files.

The first set of results obtained from a model are sometimes found to represent inadequately the actual observed concentration patterns in the region. Thus, diagnostic analyses are needed to uncover the causes of these discrepancies. For example, after the first simulation of a smoggy day in Sacramento, Reynolds et al. (1979) found that the magnitudes of the computed ozone concentrations were in general agreement with the observed values, but the predicted ozone cloud was not in the appropriate location. Suspecting a possible problem in the wind inputs, they reexamined the wind input file and the observed meteorological data. An assessment of the single available pibal sounding in light of the predicted concentration results suggested that they should not have ignored the significant directional shear indicated in the sounding. In preparing the initial wind input file, they considered wind speed changes with height but not directional changes. Upon revising the wind inputs and exercising the model, they obtained significantly improved agreement between computed and observed ozone values.

Throughout the simulation, the behavior of both primary and secondary pollutants may suggest ways of improving the model results. Comparison of the diurnal cycles of the measurements with the model predictions may indicate that the initial or boundary conditions have been set too high, that the estimates of emissions are too low, or that the atmospheric stability has been incorrectly specified. These problems can generally be discovered by reassessing the input data, and an appropriate change can then be made that improves the results. If the improvement is still insufficient to bring performance up to standard, the improved results should be reanalyzed according to the same procedure that was used initially, until the desired results are obtained.

However, results obtained by adjusting model inputs to bring calculation closer to observation should only be accepted as evidence of satisfactory model performance if the input adjustments can be justified as a consequence of independent data (e.g., wind data not used initially, remeasurement of emissions composition, etc). Simulations based on the adjustment of some input within its range of error (to improve fidelity) should be accepted only if the adjustment is general (say a reaction rate constant) and should not be used in model evaluation if the adjustment is day specific (say a wind speed or a boundary concentration).

If model results are found to be inadequate for a particular day, then selection of another day may yield improved model performance. The available data base may not be suitable for describing complex meteorological phenomena; a day with simpler meteorological conditions may be better suited for testing the model's performance.

In the development of any photochemical model, many approximations are invoked to derive a "workable" computational procedure. Some of these approximations may degrade model performance. For a particular combination of inputs, a set of approximations that under most conditions does not significantly affect the results, may generate errors that are unacceptable. If this situation can be identified, it may be possible to correct it. However, making changes in the computer code requires a thorough understanding of that code. Most users will not have enough experience with models to make code changes, so if a model deficiency is discovered the model developer should be consulted. Another possibility would be to consider the usage of a potentially more suitable model, if available, though time and resource constraints may preclude this alternative.

In summary, the diagnosis of model performance problems is generally an iterative procedure: A set of results is analyzed to determine the likely causes of the discrepancies, the inputs of the model are revised to reflect better the actual physical and chemical atmospheric phenomena, and then the model is exercised again.

To carry out the appropriate diagnostic analysis, the study team should have a thorough knowledge of the model's formulation, the computer codes, and the pertinent atmospheric and emissions phenomena, as well as experience in applying this knowledge to the resolution of model performance problems. Without these capabilities, model users are likely to pursue inappropriate directions leading to inefficiencies that will raise project costs and delay completion of a project. If the users do not possess the requisite skills, then appropriate EPA modeling experts or private consultants should be contacted to aid in the diagnosis.

It is important that model inputs be modified within their range of uncertainty. Since discrepancies between computed and measured values can result from errors in both the model's formulation and its inputs, one should not attempt to cover up formulation inadequacies through selective use of model inputs. Inadequate model performance is an important finding, possibly indicating a need to further refine one or more components of the model. Alternatively, inadequate model performance may suggest that the atmospheric phenomena that actually occurred on the simulation day cannot be sufficiently characterized by the available observational data.

Before the model is used in the application study, satisfactory performance should be demonstrated. In some cases it may not be possible to verify the model with the available data base and resources, yet circumstances may dictate that, even though unverified, the model must be used anyway. For example, the analysis may be subject to a deadline imposed by governmental regulations. Perhaps the model represents the current state of the art and no improvements are possible without a major research effort. Whatever the reason, if the model is used, the previous analyses may pinpoint some of the deficiencies in the model that caused it to fail to meet the required performance standards. These deficiencies should be fully detailed in the account of its use.

The use of "correction factors" to calibrate model results should be carefully considered and reviewed by all entities participating in or reviewing the plan. It is unclear that such calibration factors can be validly applied to subsequent control strategy results. Thus, the stress should be placed on discovery and correction of discrepancies between computed and measured values, not on calibration. Additional discussion of this issue is contained in Section 6.

EXAMPLES OF PREVIOUS MODEL EVALUATION STUDIES

Some of the most extensive photochemical model performance evaluation studies have been carried out using the Livermore Regional Air Quality (LIRAQ) Model and the SAI Airshed Model. This section provides a brief overview of these model performance evaluation efforts.

The LIRAQ Model

The LIRAQ Model was evaluated by researchers at the Lawrence Livermore Laboratory (Duewer, MacCracken, and Walton, 1978), who applied it to the San Francisco Bay Area, the region for which it was developed. Determination of the modeling region and the date and time to be simulated were based on the types of future studies that were expected to use the model. For applicability to selection of future control strategies, the modeling region would have to include all the major source areas. Since the photochemical version of LIRAQ can have at most 20 x 20 grid squares, the size of the grid squares was set at 5 km x 5 km so that the main population and industrial areas would be included while maintaining reasonable spatial resolution.

The simulation period chosen was sunrise to sunset on two days in 1973: 26 July and 20 August. The former was one of the highest oxidant days of the year, whereas the latter had relatively clean air. The period from sunrise to sunset covers the typical cycle of atmospheric events, starting with the morning emissions of primary pollutants, proceeding with the buildup of secondary pollutants, and concluding with the cessation of photochemical reaction processes when the sun sets. The 1-hour-average concentrations were calculated for comparison with the station measurements.

The LIRAQ model evaluation study was coordinated with the Bay Area Air Quality Management District (BAAQMD, formerly the BAAPCD) and other local government agencies, which were responsible for collecting most of the input data. Section 3 briefly describes how the data were collected and used to prepare the input files. Data collection and data preparation were coordinated so that the input file programs were designed to use all the available data and not to require information that was not available. This feature would not apply if LIRAQ were used for another study region.

Since LIRAQ was originally developed for the San Francisco Bay Area, model adaptation to the study area was carried out as part of the developmental efforts. The only input variable that was adjusted as part of the model evaluation was the boundary conditions. The almost complete lack of data for specifying the boundary conditions motivated a sensitivity study of the effect of the boundary conditions on model performance. In that sensitivity analysis, three sets of boundary conditions were used for each day. The results showed that the effect of the boundary conditions is significant; one set produced better model performance in terms of a variety of statistical measures.

The simulation results were compared with the station measurements of NO, NO2, total hydrocarbons, CO, and O3 (MacCracken and Sauter, 1975; Duewer, MacCracken, and Walton, 1978). Examples of the types of analyses performed are shown in Figures 4-1 and 4-2. In the plots for each station, the qualitative measure of model performance is easier to interpret than some of the statistical measures, but the quantitative measures are necessary for comparison of different evaluation studies. The scatter plots of the mean and maximum concentrations at each station provide another indicator of model performance.

Since the first evaluation study of LIRAQ, the model has been improved, particularly by the inclusion of an updated chemical mechanism. This improved model has been evaluated using the same data base used in the first effort. Subsequently, the model has been used to determine possible control strategies for meeting the NAAQS as part of the revisions being proposed to the State Implementation Plan (SIP) for the San Francisco Bay Area.

The SAI Airshed Model

The SAI Airshed Model has been applied to five U.S. cities, including Denver, Los Angeles, Las Vegas, St. Louis, and Sacramento. Table 4-2 summarizes the number of days simulated for each city and provides a list of relevant references. The discussion below is limited to the Denver evaluation effort.

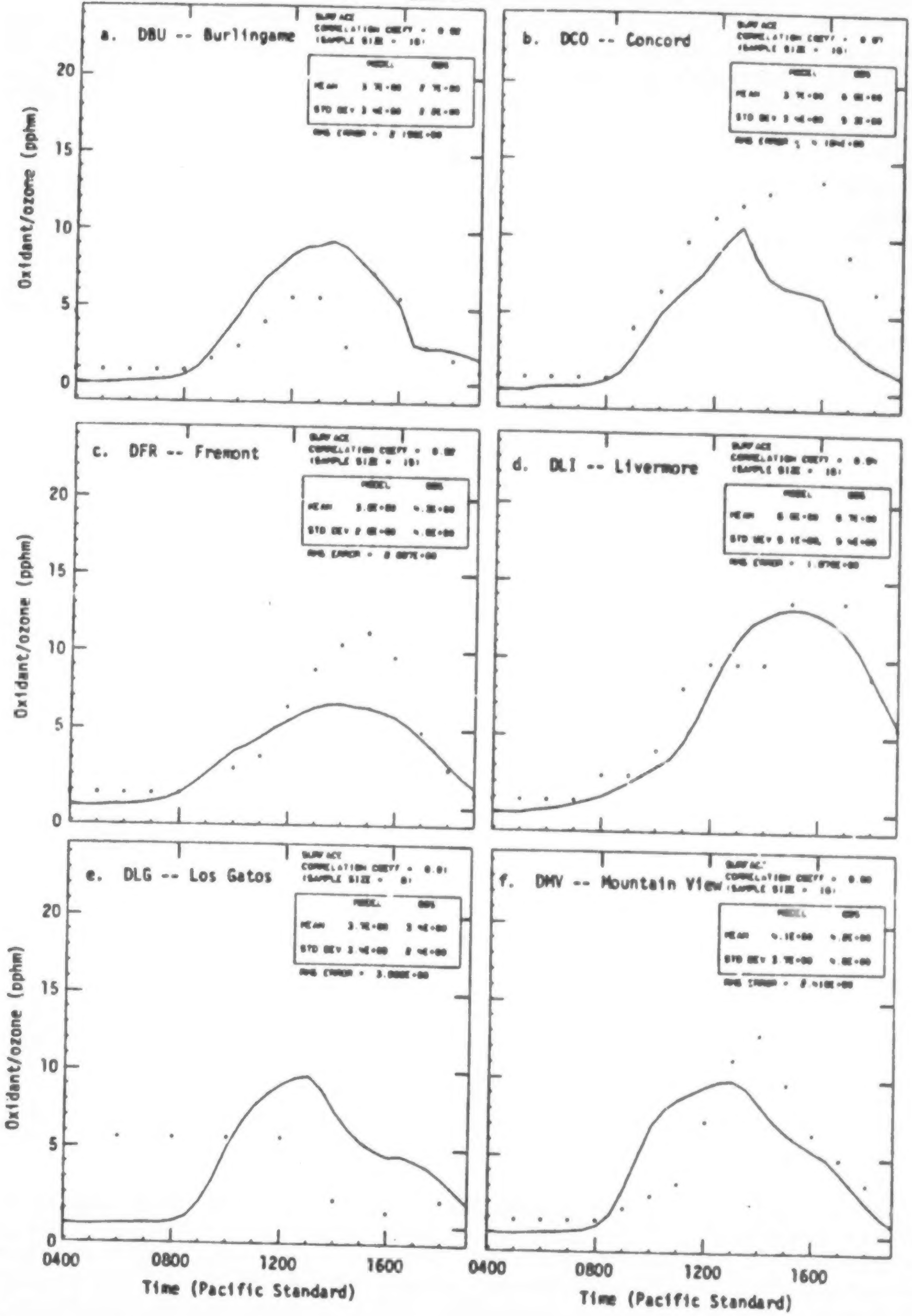
Denver Evaluation Studies. Three studies of the air quality in the Denver metropolitan area have been carried out using different versions of the SAI Airshed Model. This discussion is limited to the investigation that was done for Region VIII of the EPA (Anderson et al., 1977). That study evaluated model performance and analyzed the sensitivity of 1985 and 2000 air quality to different growth and transportation scenarios using data collected by the Colorado Departments of Health and Highways.

The daylight hours of three days were simulated for the evaluation study: 29 July 1975, 28 July 1976, and 3 August 1976. To collect additional meteorological data, a special field study was conducted on 29 July 1975. No additional data were collected for the other two days.

The data that are routinely collected in the Denver area include wind speed and direction at 10 to 15 stations, two radiosondes per day at Stapleton International Airport, and air quality measurements at about 10 locations. The emissions were calculated from data collected by the Air Pollution Control Division of the Colorado Health Department and from a highway model developed by the Colorado Highway Department. Section 3 describes the data gathering efforts in more detail.

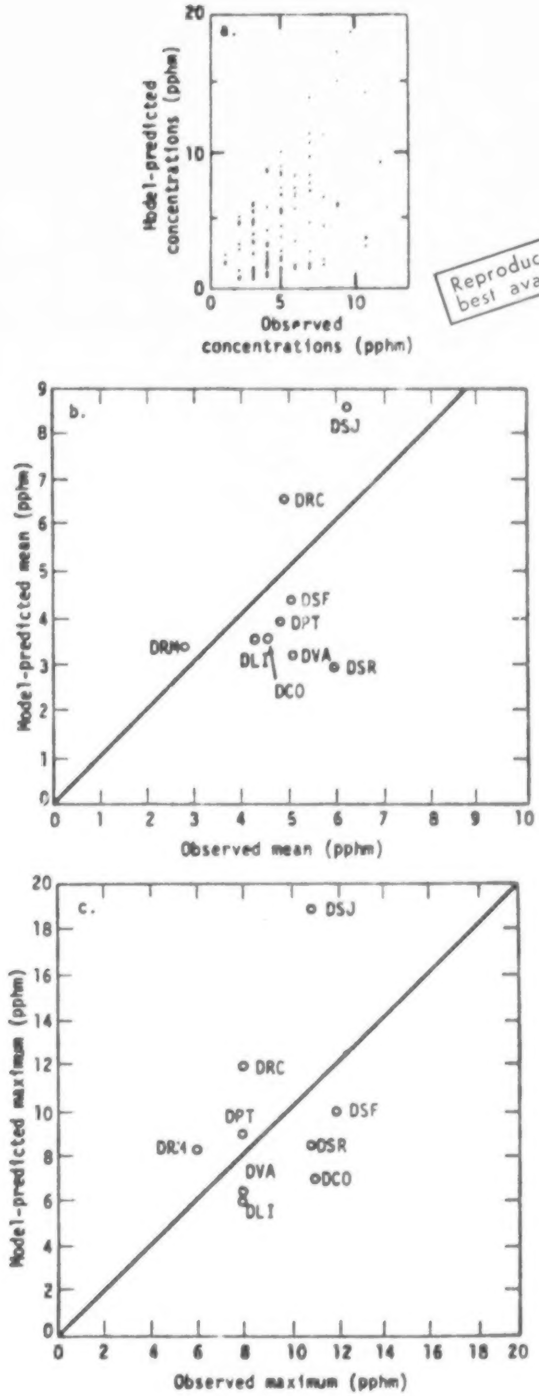
Unlike the LIRAQ evaluation application, the SAI Airshed Model had to be modified specifically for the Denver application, since the model was applied first to the Los Angeles air basin. These modifications included relatively straightforward changes to the DIMENSION and FORMAT statements in the computer programs. The latest version of the SAI

Reproduced from
best available copy.



Source: Duewer, MacCracken, and Walton (1978).

FIGURE 4-1 EXAMPLES FROM THE COMPARISON OF AVAILABLE STATION OBSERVATIONS OF OXIDANT WITH LIRAQ CALCULATED TIME HISTORIES OF OZONE



Reproduced from best available copy.

Source: Duewer, MacCracken, and Walton (1978).

FIGURE 4-2 SAMPLE SCATTER DIAGRAMS OF OBSERVED VERSUS CALCULATED HOURLY AVERAGE NO_2 CONCENTRATIONS, OBSERVED VERSUS CALCULATED STATION MEAN CONCENTRATIONS, AND OBSERVED VERSUS CALCULATED STATION MAXIMUM CONCENTRATIONS FROM THE LIRAQ SIMULATIONS

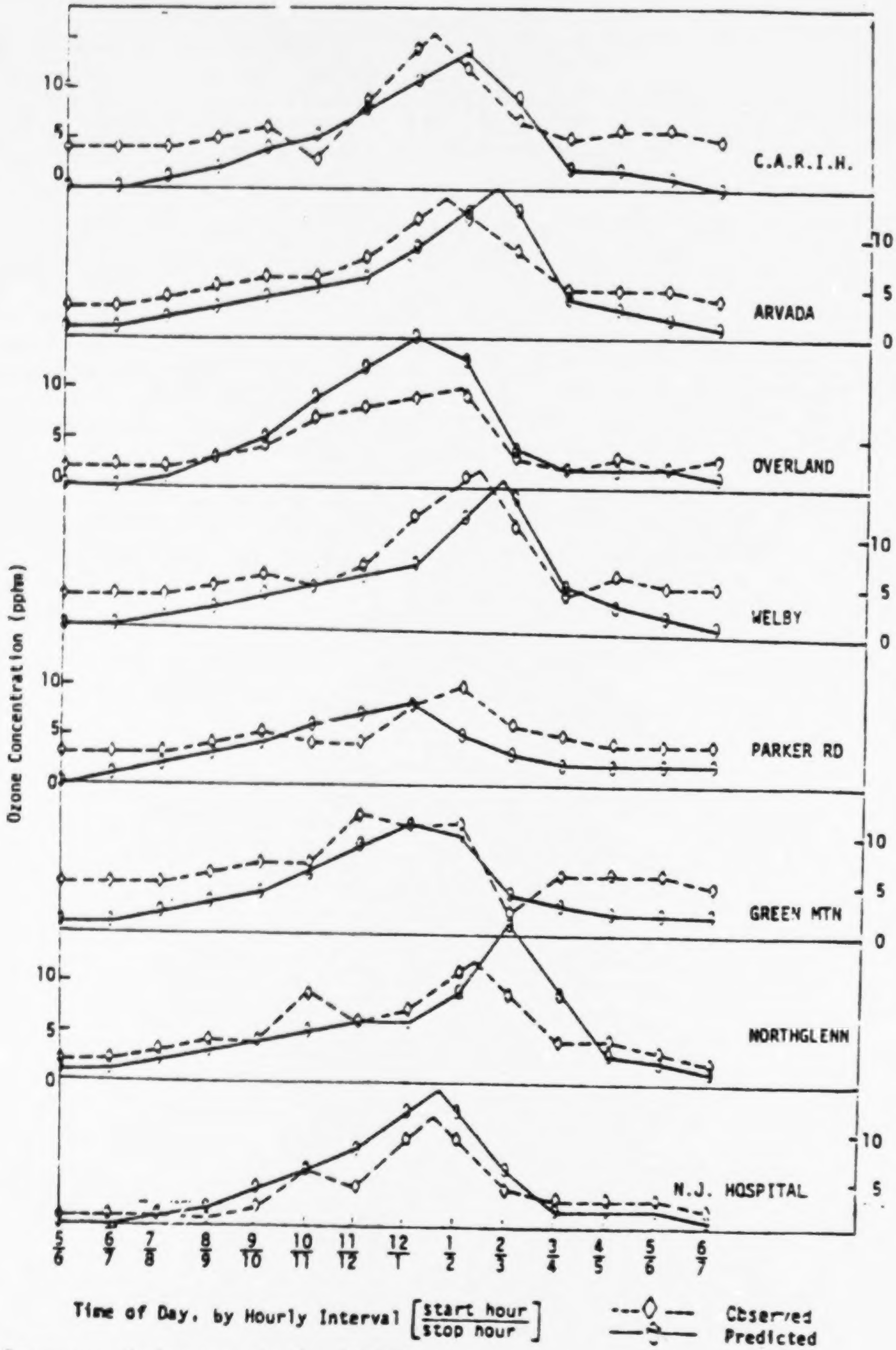
Table 4.2. Summary of SAI Experience in Evaluating the Performance of the SAI Airshed Model

City	No. of Days Simulated	Sponsor	Model Version	Reference
Los Angeles	6	EPA	1971, 1973	Roth et al. (1971); Reynolds et al. (1973)
	1	Southern California Edison	1976	Tesche and Burton (1978); Tesche, Burton, and Mirabella (1979)
	2	DOT	1976	Reynolds et al. (1979)
	2	EPA	1976	Tesche and Pollack (1978)
Denver	2	Colorado Division of Highways	1976 (Denver Model)	Donnelly (1978)
	3	EPA Region VIII	1976 (Denver Model)	Anderson et al. (1977)
	2	DOT	1976	Reynolds et al. (1979)
Las Vegas	6	EPA	1973	Liu et al. (1977)
St. Louis	3	EPA	1978	Reid et al. (1979)
Sacramento	2	DOT	1978	Reynolds et al. (1979

Airshed Model can be applied to almost any region without such modifications.

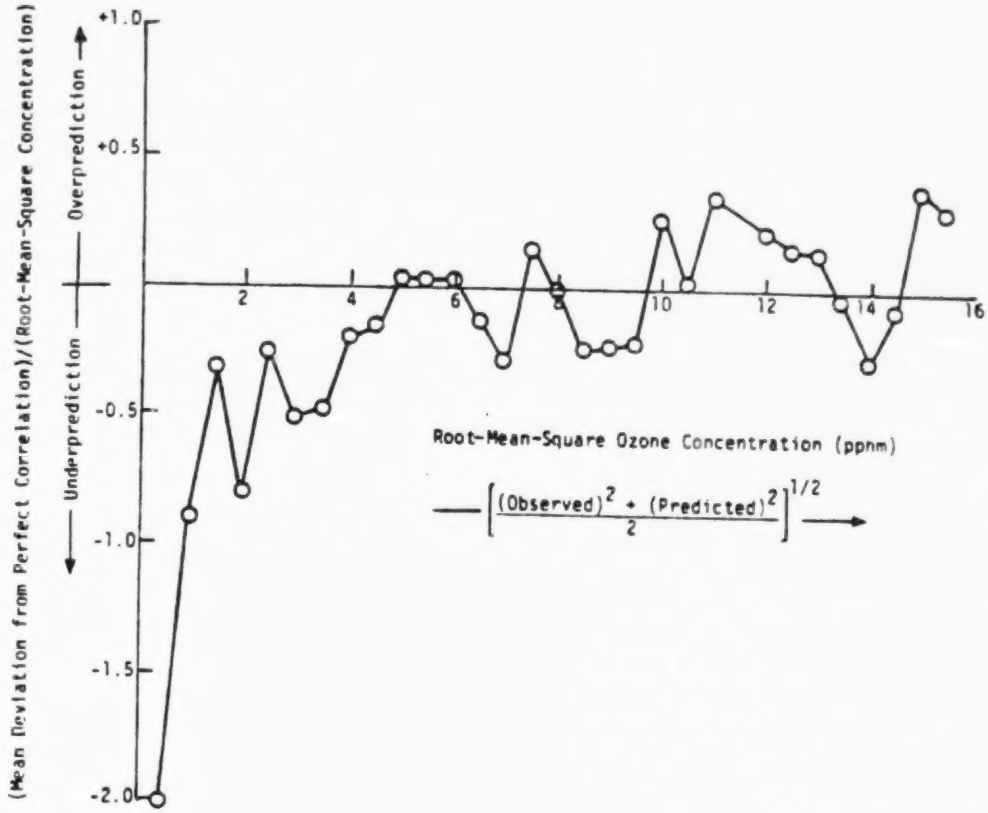
At the time of the Denver project, the Carbon-Bond Mechanism that is now incorporated in the SAI Airshed Model had just been developed. Since the Carbon-Bond Mechanism had not previously been used in the model, the simulation program and the preprocessor programs were adapted at that time to handle this chemistry routine and the species associated with it.

As is common in most applications, almost no data were available for specifying the boundary concentrations. Since the Denver modeling area had few upwind sources, the boundary concentrations were set to relatively low background levels. Once the program "bugs" and the keypunching errors were identified and removed, the model calculations agreed satisfactorily with the actual observations. Examples from some of the analyses performed for this study are given in Figures 4-3 and 4-4. Figure 4-3 shows the calculated and measured ozone concentrations for each station as a function of time. This information is the basis for all of the other statistical analyses. Figure 4-4 presents a way of estimating the model bias; it indicates that the model tends to underpredict at low concentrations, but the results show almost no bias at higher concentrations. Further analyses showed that discrepancies in the model calculations are comparable to the expected error in the observations due to measurement errors.



Source: Anderson et al. (1977).

FIGURE 4-3 EXAMPLES OF OBSERVED AND SAI AIRSHED MODEL PREDICTED HOURLY OZONE CONCENTRATIONS (pphm) AT VARIOUS STATIONS IN DENVER ON 28 JULY 1976



Source: Anderson et al. (1977).

FIGURE 4-4 ESTIMATE OF BIAS IN SAI AIRSHED MODEL PREDICTIONS AS A FUNCTION OF OZONE CONCENTRATION

REFERENCES

Anderson, G. E., et al. (1977), "Air Quality in the Denver Metropolitan Region: 1974-2000," EPA-908/1-77-002, Systems Applications, Incorporated, San Rafael, California.

Donnelly, D. E. (1978), "Oxidant Model Applications: Denver," 57th Annual Transportation Research Board Meeting, January, 1978, Washington, D. C.

Duewer, W. H., M. C. MacCracken, and J. J. Walton (1978), "The Livermore Regional Air Quality Model: II. Verification and Sample Application in the San Francisco Bay Area," J. Appl. Meteorol., Vol. 17, No. 3, pp. 273-311.

Environmental Protection Agency [EPA] (1976), "Quality Assurance Handbook for Air Pollution Measurement," EPA-600/9-76-005, Research Triangle Park, North Carolina.

Eschenroeder, A. Q., J. R. Martinez, and R. A. Nordsieck (1972), "Evaluation of a Diffusion Model for Photochemical Smog Simulation," CR-1-273, EPA-R4-73-012a, General Research Corporation, Santa Barbara, California.

Hayes, S. R. (1979), "Performance Measures and Standards for Air Quality Simulation Models," EF78-93R, Systems Applications, Incorporated, San Rafael, California.

Hillyer, M. J., S. D. Reynolds, and P. M. Roth (1979), "Procedures for Evaluating the Performance of Air Quality Simulation Models," EF79-25, Systems Applications, Incorporated, San Rafael, California.

Hilst, G. R. (1978), "Plume Model Validation," EA-917-SY, Workshop WS-78-99, Electric Power Research Institute, Palo Alto, California.

Houghland, E. S. and N. T. Stephens (1976), "Air Pollutant Monitor Siting by Analytical Techniques," J. Air. Pollut. Control Assoc., Vol. 26, p. 51.

Koch, R. C., and S. D. Thayer (1971), "Validation and Sensitivity Analysis of the Gaussian Plume Multiple-Source Urban Diffusion Model," EF-60, GEOMET, Incorporated, Gaithersburg, Maryland.

Lamb, R. G. and J. H. Seinfeld (1973), "Mathematical Modeling of Urban Air Pollution--General Theory," Environ. Sci. Technol., Vol. 7, pp. 253-261.

Lehmann, E. L. (1975), Nonparametrics: Statistical Methods Based on Ranks, (Holden-Day Incorporated, San Francisco, California).

Lui, M. K., et al., (1976), "Continued Research in Mesoscale Air Pollution Simulation Modeling: Vol. I. Assessment of Prior Model Evaluation Studies and Analysis of Model Validity and Sensitivity," EPA 600/4-76-016A, Systems Applications, Incorporated, San Rafael, California.

MacCracken, M. C., and G. D. Sauter, eds. (1975), "Development of an Air Pollution Model for the San Francisco Bay Area," UCRL-51920, Vol. 1, Lawrence Livermore Laboratory, Livermore, California.

Ott, W. R. (1977), "Development of Criteria for Siting Air Monitoring Stations," J. Air Pollut. Control Asoc., Vol. 27, p. 543.

Reid, L. E., et al. (1979), "Adaptation of the SAI Airshed Model for Usage with the Regional Air Pollution Study (RAPS) Data Base," EPA 68-02-2429, Systems Applications, Incorporated, San Rafael, California.

Reynolds, S. D. et al. (1979), "Photochemical Modeling of Transportation Control Strategies, Vol. I. Model Development, Performance Evaluation, and Strategy Assessment," EF79-28, Systems Applications, Incorporated, San Rafael, California.

__________________ (1973), "Further Development and Validation of a Simulation Model for Estimating Ground Level Concentrations of Photochemical Pollutants," Systems Applications, Incorporated, San Rafael, California.

Roth, P. M., et al. (1971), "Development of a Simulation Model for Estimating Ground Level Concentrations of Photochemical Pollutants," 71-SAI-21, Systems Applications, Incorporated, San Rafael, California.

Seinfeld, J. H. (1977), "Current Air Quality Simulation Model Utility," Department of Chemical Engineering, California Institute of Technology, Pasadena, California.

__________________ (1972), "Optimal Location of Pollutant Monitoring Stations in an Airshed," Atmos. Environ., Vol. 6, p. 847.

Tesche, T. W., and C. S. Burton (1978), "Simulated Impact of Alternative Emissions Control Strategies on Photochemical Oxidants in Los Angeles," EF78-22R, Systems Applications, Incorporated, San Rafael, California.

Tesche, T. W., and R. I. Pollack (1978), "Evaluating Simple Oxidant Prediction Methods Using Complex Photochemical Models," EPA 68-02-2870, Systems Applications, Incorporated, San Rafael, California.

Tesche, T. W., C. S. Burton, and V. A. Mirabella (1979), "Recent Verification Studies with the SAI Urban Airshed Model in the South Coast Air Basin," Proc. of Fourth Symposium on Turbulence, Diffusion, and Air Pollution, 15-18 January 1979, Reno, Nevada.

Wayne, L. G., A. Kokin, and M. I. Weisburd (1973), "Controlled Evaluation of the Reactive Environmental Simulation Model (REM)," EPA R4-73-013a, Pacific Environmental Services, Incorporated, Santa Monica, California.

5. MODEL APPLICATIONS

Model applications to support the development of a SIP can take a variety of forms. As a minimum, the model should be applied to demonstrate that a control strategy or set of emission reductions will result in attainment of the ozone standard. This single application forms the bridge between the problem, as defined by excesses of the ambient ozone standard, and the proposed solution. The problem occurs, however, of how to anticipate the appropriate control strategy. Since there have been only two previous applications of photochemical dispersion models in support of air quality plan development, the base of available experience in this area is limited.

The approach emphasized in this report is the approach used in the San Francisco Bay Area. It consisted of three types of applications: baseline projections, emissions sensitivity analyses, and control strategy simulations. Depending on the specific issues to be addressed and resource constraints, the degree to which one or another of these applications is employed will be unique to each modeling program. For example, it may be more economical to omit or minimize the use of sensitivity tests if the control strategies to be tested are well-defined at an early stage in the process. Other approaches or refinements may be developed in the future as more experience is gained in model applications.

BASELINE PROJECTION

This application establishes the magnitude of air quality problems in future years in the absence of any additional control programs. It addresses the question of whether the ozone standard will be met at some point in the future. For example, the 1977 Clean Air Act Amendments attainment deadlines of 1982 and 1987 would be likely choices for future years to be used in the modeling analysis. If the results of the baseline runs for those years indicate that excesses of the ozone standard will still occur despite the continued application of all existing control programs, then the need for additional control programs is established. If the results indicate that attainment may be expected sometime between 1982 and 1987, then any additional controls considered could be focussed on what could be implemented in the near term. If attainment is indicated in both 1982 and 1987, then there is no need for a non-attainment plan and attention can be focussed on long-term maintenance of the ozone standard.

The Clean Air Act also requires long-term maintenance of air quality standards, and EPA has interpreted this to mean a minimum of ten years following the expected attainment date. Thus, an additional future year of interest would be further into the future, such as the year 2000. The selection of the particular year to be used in this case may depend more heavily on the availability of data on which the emission inventory projection may be based.

Preceding page blank

EMISSIONS SENSITIVITY ANALYSES

There are numerous sensitivity tests that may be conducted to answer a variety of questions. One set of tests that would be central to plan development would be to simulate the effect of various hydrocarbon and/or NOx emission reductions in order to estimate appropriate targets for controls. In this instance, it may be useful to set up a diagram of the tests as illustrated in Figure 5-1. This serves to display the potential test options for discussion. (If resources are abundant enough tests may be run such that an EKMA-type isopleth diagram could be derived which would be unique to the region being modeled.)

The principal advantage of sensitivity tests is that they may be structured to focus on whatever control issues are important in a given region. Volume II of this report describes the applicability of various selected photochemical models to a variety of strategic control issues (e.g., control of elevated vs. low-level emissions, selective control of hydrocarbon species, long range transport, effects of spatial and temporal redistribution of emissions, etc.). In each case, care should be taken in the construction of the tests, particularly in the setting of initial and boundary conditions for each run. (See Sections 3 and 6 for additional discussion of initial and boundary conditions.)

CONTROL STRATEGY SIMULATIONS

It is likely that more than one control strategy simulation will be desirable because (1) there will likely be more than one combination of control measures that could result in attainment of the ozone standard, and (2) there may be particular interest in testing the effectiveness of specific controls. The structure of these simulations is determined solely by the priorities for each, and the time and resources available at this stage of plan development. It should be kept in mind that for plan development purposes, these simulations are the culmination of all of the previous efforts to collect data, evaluate model performance, and perform baseline and sensitivity tests. Again, particular care should be taken in specifying the initial and boundary conditions for these runs.

ALTERNATIVE PROGRAM DESIGNS

One approach to designing a model application program is to emphasize the use of sensitivity analyses to assist in the structuring of control options. For example, analyses could focus on estimating the impact on ambient ozone levels of pre-specified percentages of hydrocarbon and NOx emission reduction. These percent emission reductions could be applied uniformly without regard to location, time, or reactivity of the emission. This sensitivity study would generate target percent emission reduction estimates for hydrocarbon and nitrogen oxides emissions. An emission control strategy could then be designed to result in the given percent emission reductions, and would then be subsequently tested by the model.

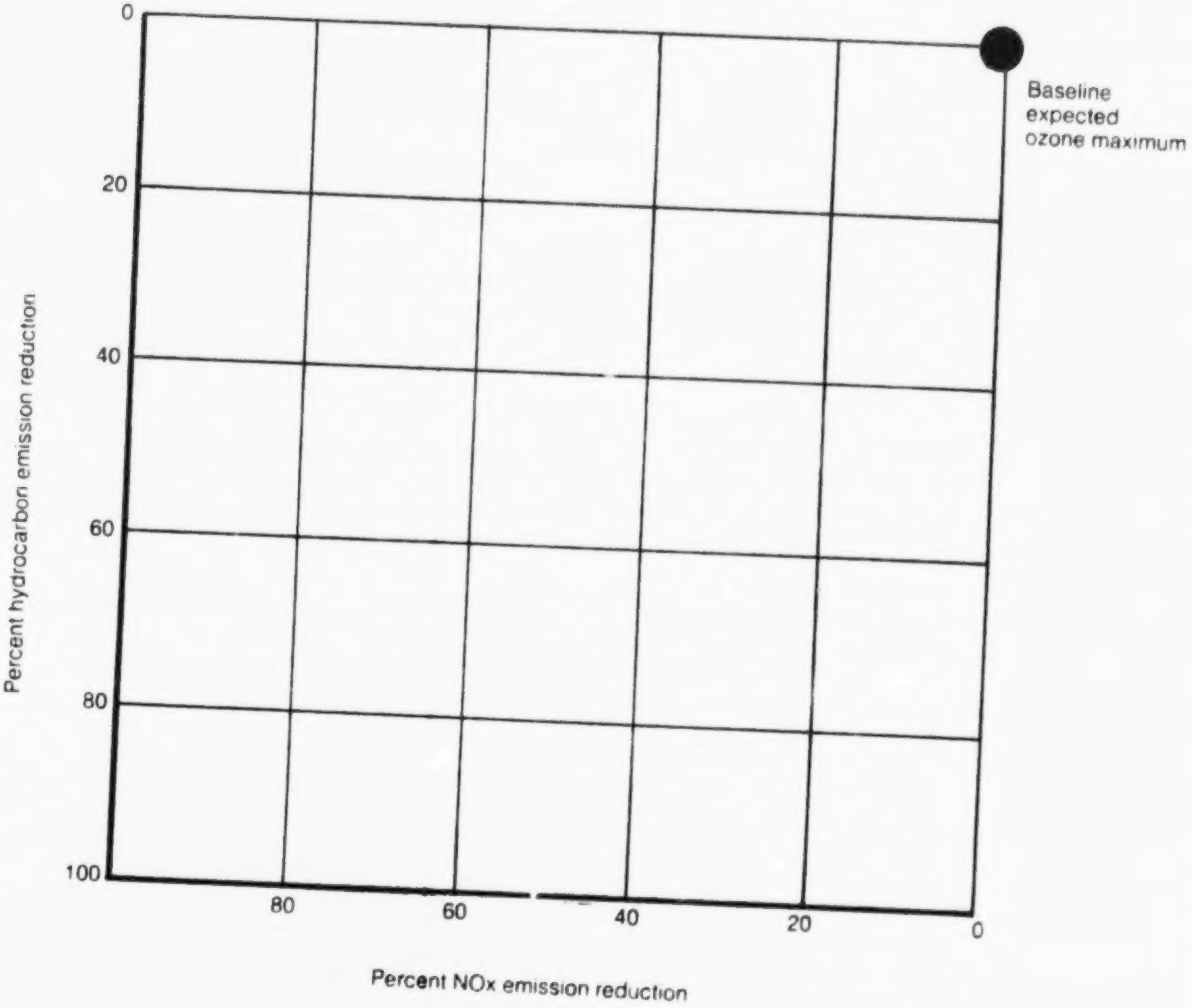


Figure 5-1

Sample diagram of potential emissions sensitivity test cases

A second approach would emphasize the direct testing of alternative control strategies. This approach would involve an iterative testing of various control strategies. For example, the first strategy tested may be a strategy which meets the emission reduction requirements estimated by EKMA. If the resulting ozone concentration is below the standard, then the next strategy to be tested would be less stringent. This process continues until an optimum strategy that meets the ambient air quality standard is derived.

The specific approach adopted in a given program will be unique to the issues and constraints of the total planning effort.

PRACTICAL ASPECTS OF DESIGNING A MODEL APPLICATION STUDY

A sample sequence of tasks for completing each of the model applications mentioned is shown in Figure 5-2. The sequence is consistent with the timing required to produce the emission inventory inputs. Baseline inventories can be prepared first, while inventories that include the effects of alternative controls require additional development effort and are therefore done last. In between, the emissions sensitivity analyses can be done using the baseline inventories, and can provide useful guidance on how stringent the control strategies should be designed to be.

The amount of time and resources devoted to each application will depend on the number of prototype days being used, the number of future years under consideration, the number of different strategies to be examined, etc. These parameters will be unique to each region.

The number of model runs required beyond those used to evaluate model performance may be estimated with the following formula:

$$N = D \times Y \times (1 + S + C) \times M$$

where

N = number of model runs
D = number of prototype days
Y = number of future years under consideration
S = number of sensitivity analysis cases
C = number of control strategy cases
1 is indicated to account for baseline runs
M = arbitrary multiplier to account for wasted runs caused by errors in input specification or other factors.

For example, if two prototype days and two future years are being used to evaluate six sensitivity cases and three control strategies, then the number of useful model runs needed is 40. If M=2, then the total number of runs becomes 80. In many cases, and especially in the case of Eulerian grid models, it may not be practical to complete this number of runs. There are a variety of ways to reduce the number of model runs:

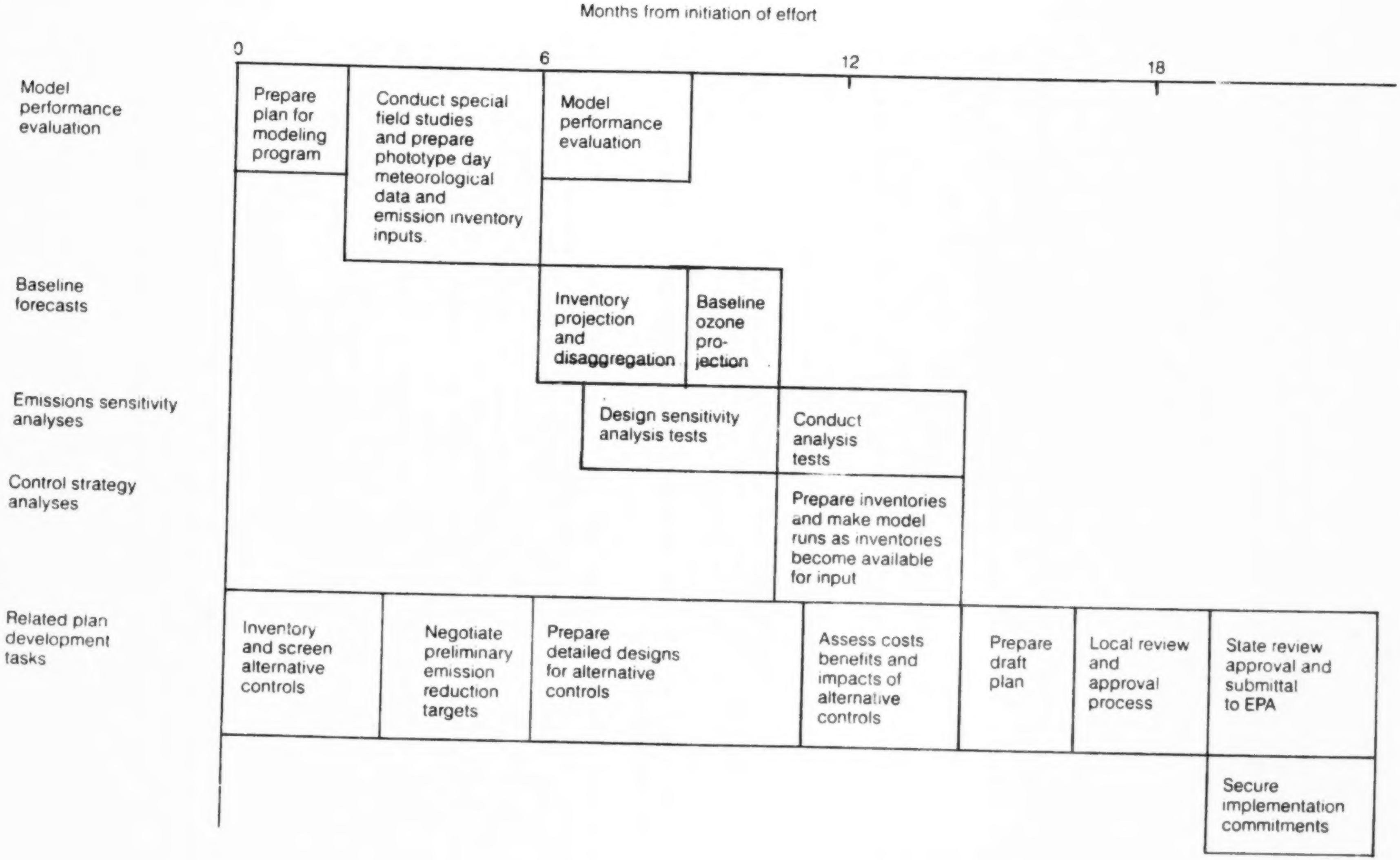

Months from initiation of effort
0
6
12
18
Model performance evaluation
Baseline forecasts
Emissions sensitivity analyses
Control strategy analyses
Related plan development tasks
Prepare plan for modeling program
Conduct special field studies and prepare phototype day meteorological data and emission inventory inputs.
Model performance evaluation
Inventory projection and disaggregation
Baseline ozone projection
Design sensitivity analysis tests
Conduct analysis tests
Prepare inventories and make model runs as inventories become available for input
Inventory and screen alternative controls
Negotiate preliminary emission reduction targets
Prepare detailed designs for alternative controls
Assess costs benefits and impacts of alternative controls
Prepare draft plan
Local review and approval process
State review approval and submittal to EPA
Secure implementation commitments

Figure 5-2

Sample schedule for model application tasks

- The sensitivity analysis tests may be restricted to a single future year and a single prototype day (in the example, six sensitivity runs would be made);
- Control strategies to be tested may be lumped together, and tests may be restricted to two prototype days (in the example, D=2, Y=2 and C=3, resulting in 12 control strategy runs);
- The EKMA model or a trajectory model may be employed to perform sensitivity analysis tests.

Care should be exercised in interpreting the results of the sensitivity tests if a different model is being employed for that purpose. Due to their unique construction, each model may provide results which appear qualitatively similar but may quantitatively differ. The ultimate usefulness of applying more than one model in the planning effort is limited due to the likelihood of discrepancies between results. Even small discrepancies of ten percent can translate into substantial cost differences in the controls required to meet the standard, thus reducing the credibility of the plan. Thus, if EKMA is used in conjunction with a photochemical dispersion model, it should only be used to produce informal initial estimates of emission control requirements.

The multiplier factor should not be underestimated, for even the most careful programmer can overlook some minute detail that could invalidate a run. In the San Francisco example to follow, over seventy runs were made of the LIRAQ model (at an average cost of $600 per run not including staff costs), but only thirty were successful. The multiplier in that case was 2.5.*

Given that there are time and resource constraints, it should be apparent that very early in the planning process an inventory of desired model runs should be prepared. The time and resource constraints should be translated into an estimate of total model runs possible once the constraints are known and the model has been selected. (As indicated in Section 2, the constraints may influence the choice of the model to be used.) The desired runs should then be ranked in priority for execution. The resulting list may change over the course of the project due to any number of factors such as computer problems, inaccurate cost estimates, slipped deadlines, unanticipated political interest in a particular control, etc., such that the need for flexibility and periodic review of the list should be anticipated.

---

*The successful runs were divided roughly equally among the baseline, sensitivity and control strategy applications described. Greater emphasis would have been placed on the control strategy simulations if not for time and budget problems encountered at the end of the planning program.

MODEL APPLICATIONS IN THE SAN FRANCISCO BAY AREA

The specific model applications made in the San Francisco Bay Area included baseline forecasts, emission sensitivity analyses, and control strategy simulations.

Baseline Forecasts

Forecasts of future ozone trends in 1985 and 2000 were made using prototype meteorology for July 26, 1973, and emission inventories projected and disaggregated for these future years. The results of these model runs are summarized in Table 5-1. A sample isopleth map is shown in Figure 5-3. The projections indicated that regional ozone levels are expected to improve between 1975 and 1985, and that the maximum level would be reduced by approximately 20%. This improvement was attributed primarily to the effect of continued implementation of the federal and California motor vehicle emission control programs. Between 1985 and 2000, due to growth in population, motor vehicles, and normal urban activities reflected in the projected emission inventories, ozone levels were projected to deteriorate to about the 1975 levels.

Emission Sensitivity Analyses

To define the emission reductions needed to meet the oxidant standard, 1985 baseline emission levels were systematically reduced and then input to the LIRAQ model. The results of the sensitivity analysis are summarized in Table 5-2 and Figure 5-4, and led to two key conclusions regarding control strategy design:

- Reduction of hydrocarbon emissions alone was more effective in reducing ozone levels than the combined reduction of hydrocarbon and nitric oxide emissions (NOx emission reductions resulted in increased ozone levels within the study area under the conditions being modeled).

- An approximate 43% reduction in hydrocarbon emissions in 1985 would result in attainment of the 0.08 ppm oxidant standard. (Subsequent analysis indicated that an approximate 27% hydrocarbon emission reduction in 1985 would result in attainment of the 0.12 ppm ozone standard.)

Control Strategy Simulations

The effectiveness of alternative control strategies was determined by applying a complete land use, transportation, emission and air quality modeling system as illustrated in Figure 5-5. Each measure in the strategy was translated into the appropriate variable and parameter values or into an adjustment of the emissions inventory. The methods for doing this are described below.

- Technological Controls. Technological controls were tested with relative ease because they did not involve significant changes in human activities. Rather, they involved the

Table 5-1. Baseline LIRAQ Projections for the San Francisco Bay Area

	1975	1985	2000
Regionwide High Hour Ozone (ppm) (9.5 kms SEE of Livermore)	.17	.13	.17
Ozone at Highest Station (Livermore)	.13	.10	.15

Table 5-2. LIRAQ Emission Sensitivity Analysis Results

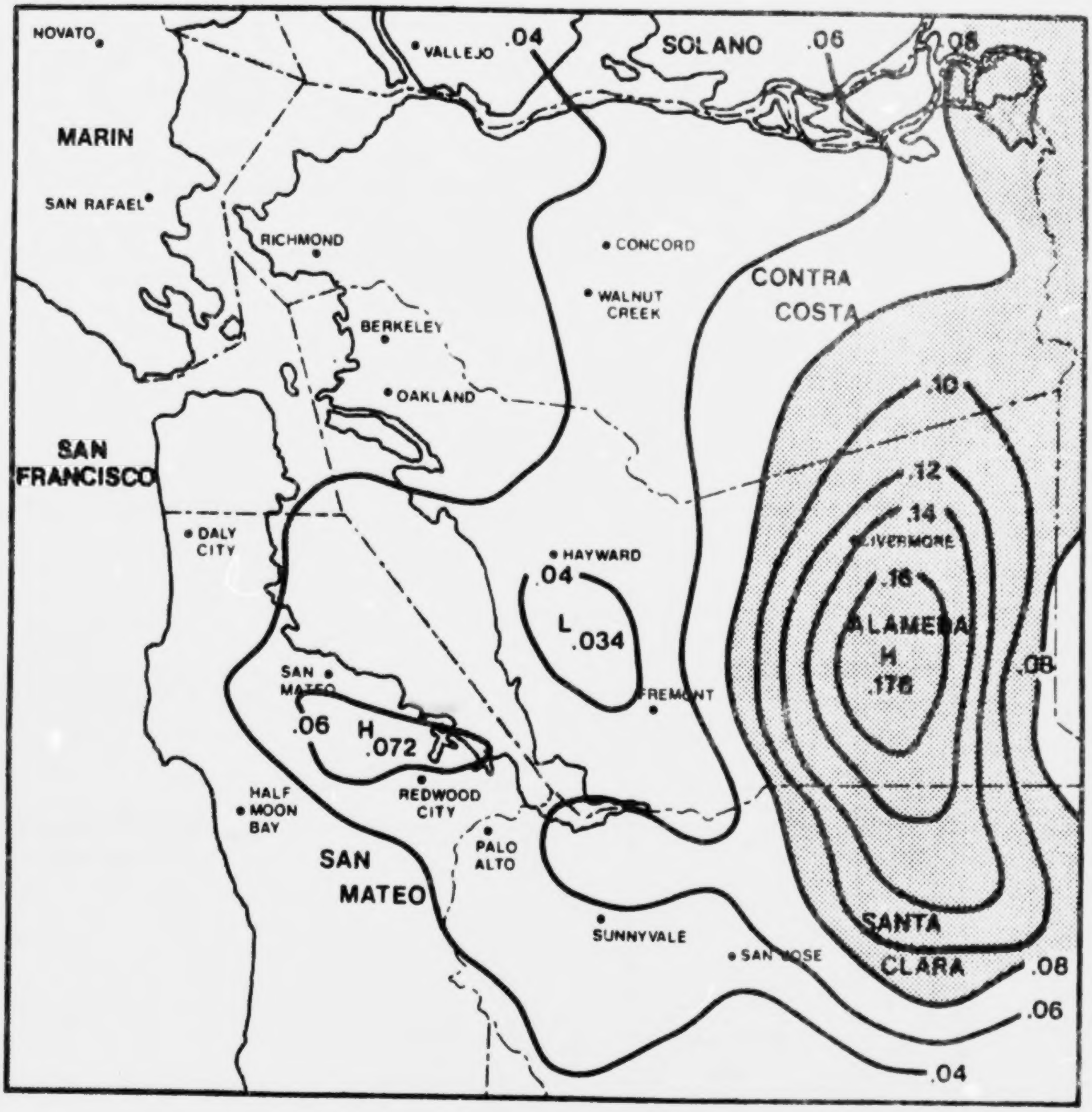
% Reduction HC	0	20	40	60	80	40	80
% Reduction NO	0	0	0	0	0	20	40
Expected worst case regionwide high hour ozone (ppm)	.19	.14	.08*	.07	.06	.11	.06

---

*This value was rounded off from an original value of .0846 ppm.

Assumptions:
1) 1985 Baseline Emission Inventory
2) July 26, 1973 Prototype Meteorology

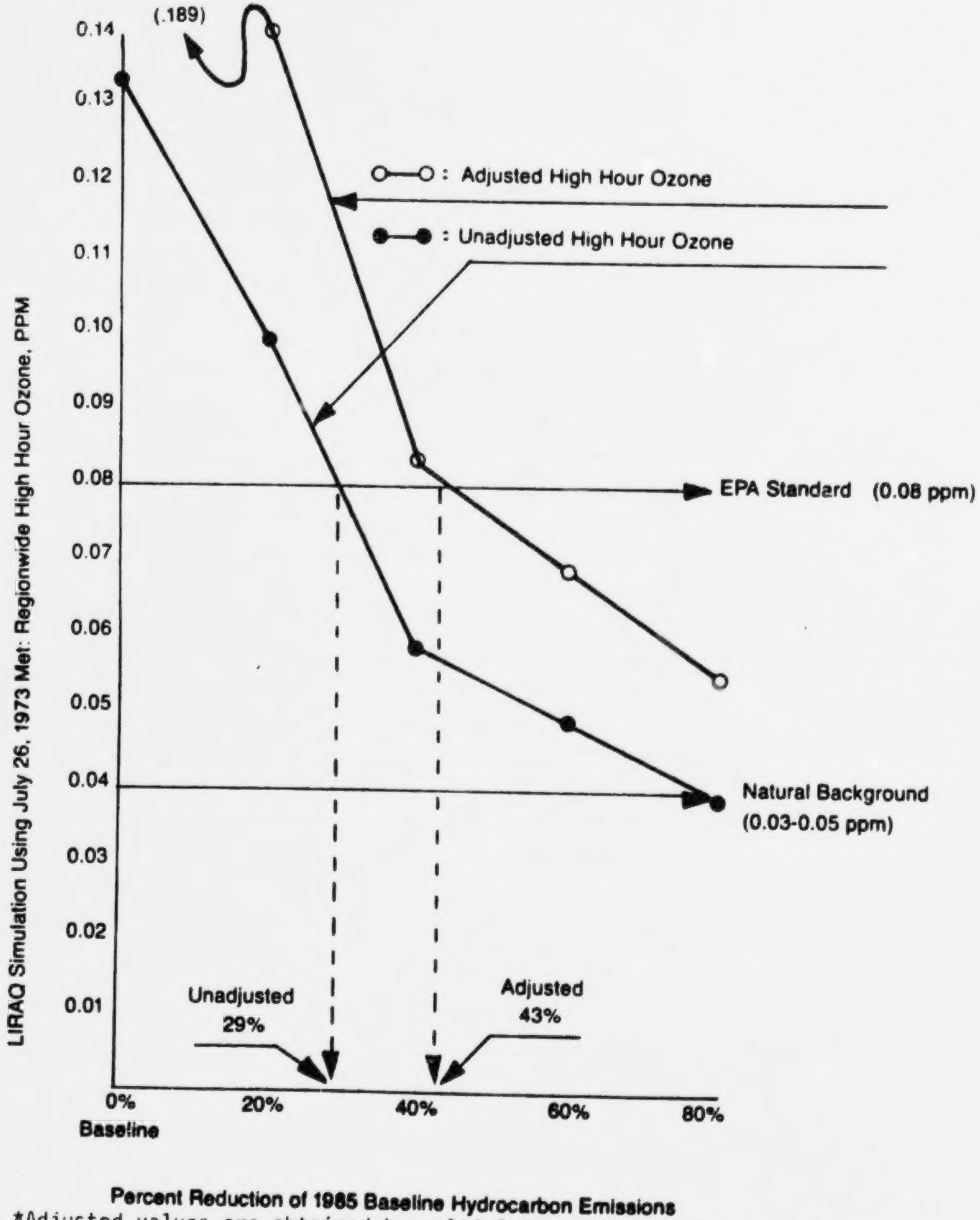
Figure 5-3. Example LIRAQ Results - 2000 Baseline Ozone Projections


NOVATO
VALLEJO
.04
SOLANO
.06
.08
MARIN
SAN RAFAEL
RICHMOND
CONCORD
WALNUT CREEK
CONTRA COSTA
BERKELEY
OAKLAND
.10
SAN FRANCISCO
.12
.14
DALY CITY
HAYWARD
LIVERMORE
.04
.16
L
.034
ALAMEDA
H
.176
.08
SAN MATEO
FREMONT
.06
H
.072
HALF MOON BAY
REDWOOD CITY
PALO ALTO
SAN MATEO
SUNNYVALE
SANTA CLARA
SAN JOSE
.08
.06
.04

July 26, 1973 Prototype Meteorology (1500 Hours PST)

Figure 5.4

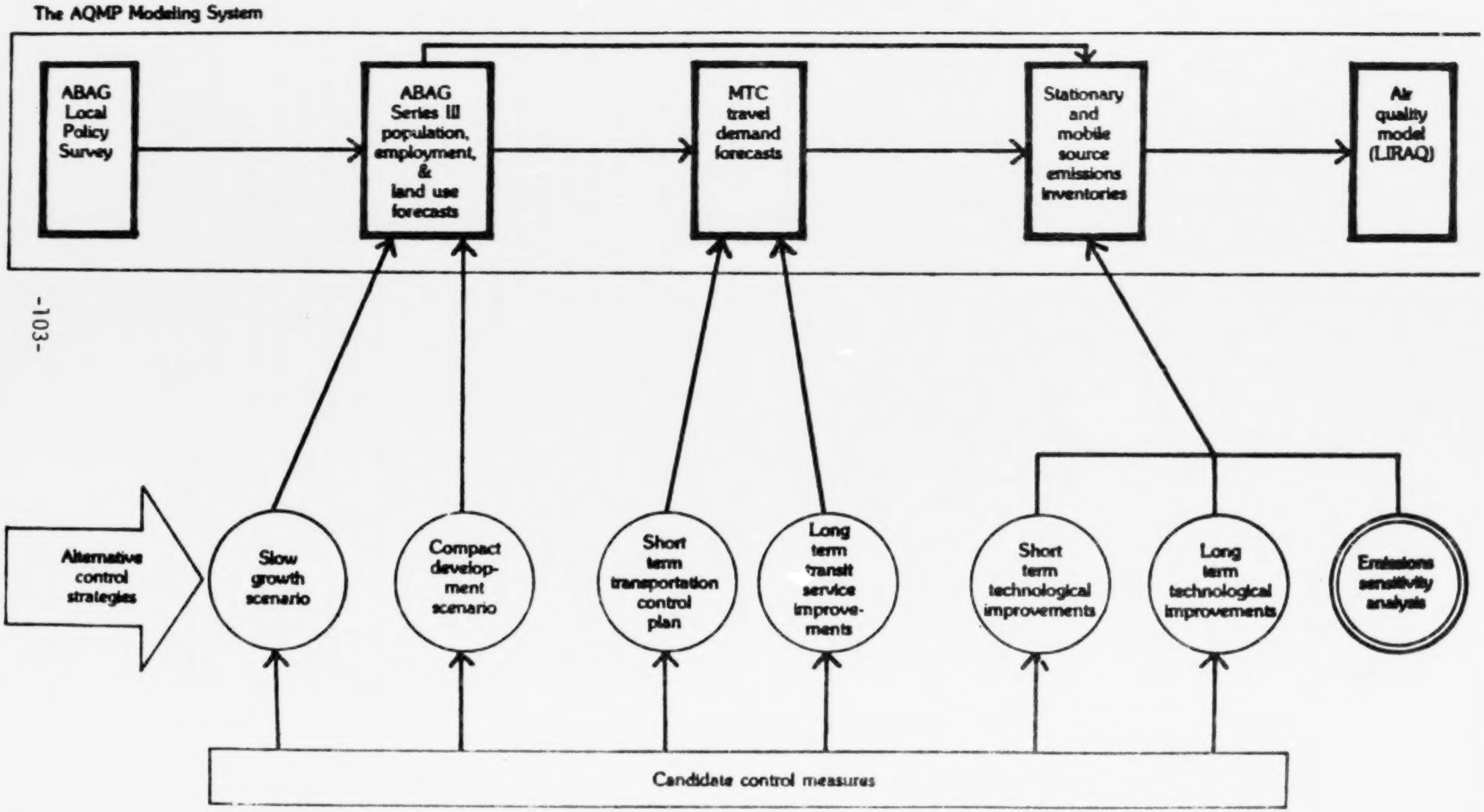
PLOTS OF UNADJUSTED AND ADJUSTED* REGIONWIDE HIGH HOUR OZONE AS A FUNCTION OF % REDUCTIONS OF 1985 HC EMISSIONS



*Adjusted values are obtained by multiplying model results by factors that account for worst-case scenarios and imperfect model performance. Such adjustments are discussed in Chapter 6.

Figure 5.5.

Control strategy testing with the AQMP Modeling System


The AQMP Modeling System
ABAG Local Policy Survey
ABAG Series III population, employment, & land use forecasts
MTC travel demand forecasts
Stationary and mobile source emissions inventories
Air quality model (LIRAQ)
Alternative control strategies
Slow growth scenario
Compact develop-ment scenario
Short term transportation control plan
Long term transit service improve-ments
Short term technological improvements
Long term technological improvements
Emissions sensitivity analysis
Candidate control measures

implementation of improved techniques for reducing the pollutant emissions resulting from normal human activities. Such emission reductions were accounted for by applying a percentage reduction factor to the "emission factors" used in the emissions models. For example, requiring even more stringent control of motor vehicle emissions than currently required was reflected in future motor vehicle emission factors. This served as input to the emissions calculations which subsequently were input to the LIRAQ model. Regulations for controlling volatile organic compounds or for implementing combustion modifications to reduce nitrogen oxide emissions from small industrial and utility boilers were handled similarly.

o Transportation Controls. Transportation controls were tested through the travel demand modeling system. Depending on the specific nature of the controls, different approaches to simulating their effects were taken. For example, the effects of a general regionwide improvement in transit service were tested by changing the transit travel time or "wait time" in the modal split model. This produced an estimate of the percent of total trips diverted to transit and produced a net decrease in highway network traffic. Testing service improvements in specific areas involved changing the transit network to reflect the improvements. Cost incentives/disincentives such as a gasoline tax or increased parking costs were simulated in the modal split model.

o Land Use Controls. The effectiveness of individual land use control mechanisms could not be tested by the forecasting system in a straightforward manner. What could be tested were the ultimate objectives of land use control measures. For example, one policy goal of land use control for improving air quality was to halt the outward spread of the metropolitan area boundaries and redirect future growth into existing urbanized portions of the region. The effectiveness of specific mechanisms or tools which might be employed to accomplish this result (e.g., tax incentives/disincentives, public facility restrictions, changes in general plans and/or zoning ordinances) could not be tested by the forecasting system. Instead, the system was used to test the effect of accomplishing that "compact development" policy goal on regional air quality. The land use policy goal in effect became an assumption for a subsequent reiteration of the ABAG forecasts. The results of these forecasts were then fed through the modeling sequence to produce estimates of resulting air quality. The information thus obtained was used to evaluate the air quality effects of a more compact development pattern in the region.

Land use controls or objectives were the most difficult and time-consuming to forecast. This was due not only to the difficulties in developing clear statements of the policy goals, but also the fact that changes in the ABAG demographic forecasts necessitated additional runs of the subsequent travel demand, emissions, and air quality models.

A summary of the control strategies tested with the modeling system is presented in Table 5-3. The schematic flow diagram of the modeling system and how alternative strategies or sensitivity analyses were input are shown in Figure 5.5. The main results of the strategy analysis are summarized in Table 5-4. The table indicates that substantial improvements in air quality can be made through the use of source control technology. The transportation and land use management strategy, although relatively ineffective in the short term, is shown to become increasingly effective with time.

MODEL APPLICATIONS IN THE DENVER METROPOLITAN AREA

Model applications in the Denver Metropolitan area also included baseline forecasts and emissions sensitivity analyses. No actual control strategy simulations have as yet been made using the SAI urban airshed model.

Baseline Forecasts

Forecasts of future ozone trends were also made for 1985 and 2000. Prototype meteorological conditions for two separate days were used in the baseline forecast. The results of these forecasts are summarized in Figure 5-6. Hourly ozone concentrations predicted for grid squares containing monitoring stations are plotted for each year. The area between results for successive years is shaded, with hatching between 1976 and 1985, and cross-hatching between 1985 and 2000. In general, reductions are greater in the 1976 to 1985 period than in the 1985 to 2000 period. Reductions of peak levels averaged 44% in the earlier period and 23% in the later period. (Additional baseline forecasts were subsequently made for 1982 and 1987.)

Sensitivity Analyses

A series of sensitivity tests were made of the effect of spatial redistribution of emissions in the Denver region. Emissions were reduced in various geographic sectors of the region to determine the impact of emissions from each area on ozone levels. The main conclusion was that perturbations in emission rates of 25 percent in different communities had no effect on region-wide ozone concentrations.

Two additional sensitivity analysis cases were run which simulated (1) a 30 percent reduction in both hydrocarbon and NOx emissions and (2) a 65 percent increase in NOx emissions. The conclusions from these tests were:

- A 30 percent reduction in all emissions resulted in only a 15 percent drop in the predicted ozone concentration.

Table 5-3. Summary of Control Strategies Tested in the San Francisco Bay Area

MAXIMUM TECHNOLOGY STRATEGY	TRANSPORTATION AND LAND USE MANAGEMENT STRATEGY	COMPREHENSIVE STRATEGY
o Use paints and other coatings that are water based and/or have a high solids content. o Use closed systems for storage and transfer of organic liquids. o Use best available control technology (BACT) on new and existing sources of hydrocarbon emissions. o Adopt more stringent vehicle (light and heavy duty) exhaust emission standards. o Implement mandatory annual inspection and maintenance program for light and heavy duty vehicles. o Require exhaust control devices on existing heavy duty gasoline trucks.	o Increase tolls on bridges. o Implement regional parking strategy to discourage private auto use and encourage high-occupancy auto use - parking tax - parking fees at large shopping centers - preferential parking for carpools, vanpools o Provide additional transit service. o Increase bus/carpool lanes and ramp metering. o Implement an auto control zone in San Francisco central business district to reduce traffic. o Provide more ride sharing services such as jitneys and vanpools. o Develop more extensive bicycle systems. o Achieve more compact development throughout the region by the year 2000.	o By 1985, the comprehensive strategy includes: all of the technological control measures except for more stringent vehicle exhaust emission standards; and all of the land use/transportation measures. The effects of compact development were not included in the analysis for 1985 since the short time frame was insufficient for achieving significant results. o By 2000, the comprehensive strategy includes: all of the technological control measures except for the exhaust control devices on existing heavy duty gasoline trucks (this measure provides short term benefits only); and all of the land use/transportation measures.

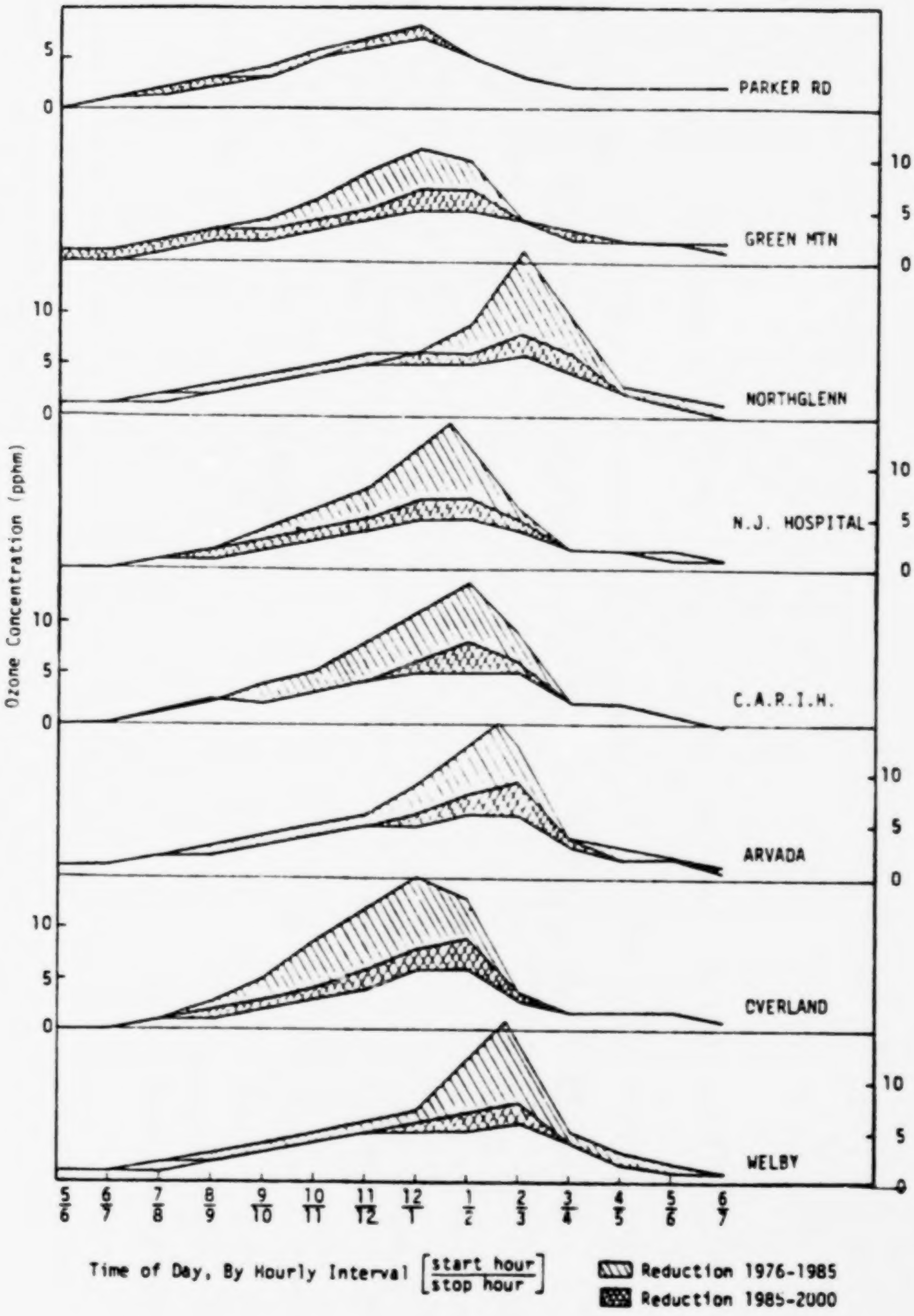
Reproduced from
best available copy.

Table 5-4. Effectiveness of Alternative Control Strategies for the San Francisco Bay Area

Strategy	1985			2000		
	Hydrocarbon Emission Reduction Potential	Estimated Regionwide High Hour Oxidant Level (ppm) [a]	Estimated No. of Annual Violations of the 1-Hour .08 ppm Federal Oxidant Standard	Hydrocarbon Emission Reduction Potential	Estimated Regionwide High Hour Oxidant Level (ppm) [a]	Estimated No. of Annual Violations of the 1-Hour .08 ppm Federal Oxidant Standard
Baseline (do-nothing)*	(797 tons/day) emitted	.19ppm	130	(1,058 tons/day) emitted	.24ppm	275
Maximum Technology	- 280 tons/day	.10ppm	3	- 441 tons/day	.13ppm	16
Transportation and Land Use Management	- 7 tons/day	not estimated	-	- 84 tons/day with slow growth	.23ppm	220
Comprehensive Strategy*	- 286 tons/day	.10ppm	3	- 513 tons/day with slow growth	.12ppm	11

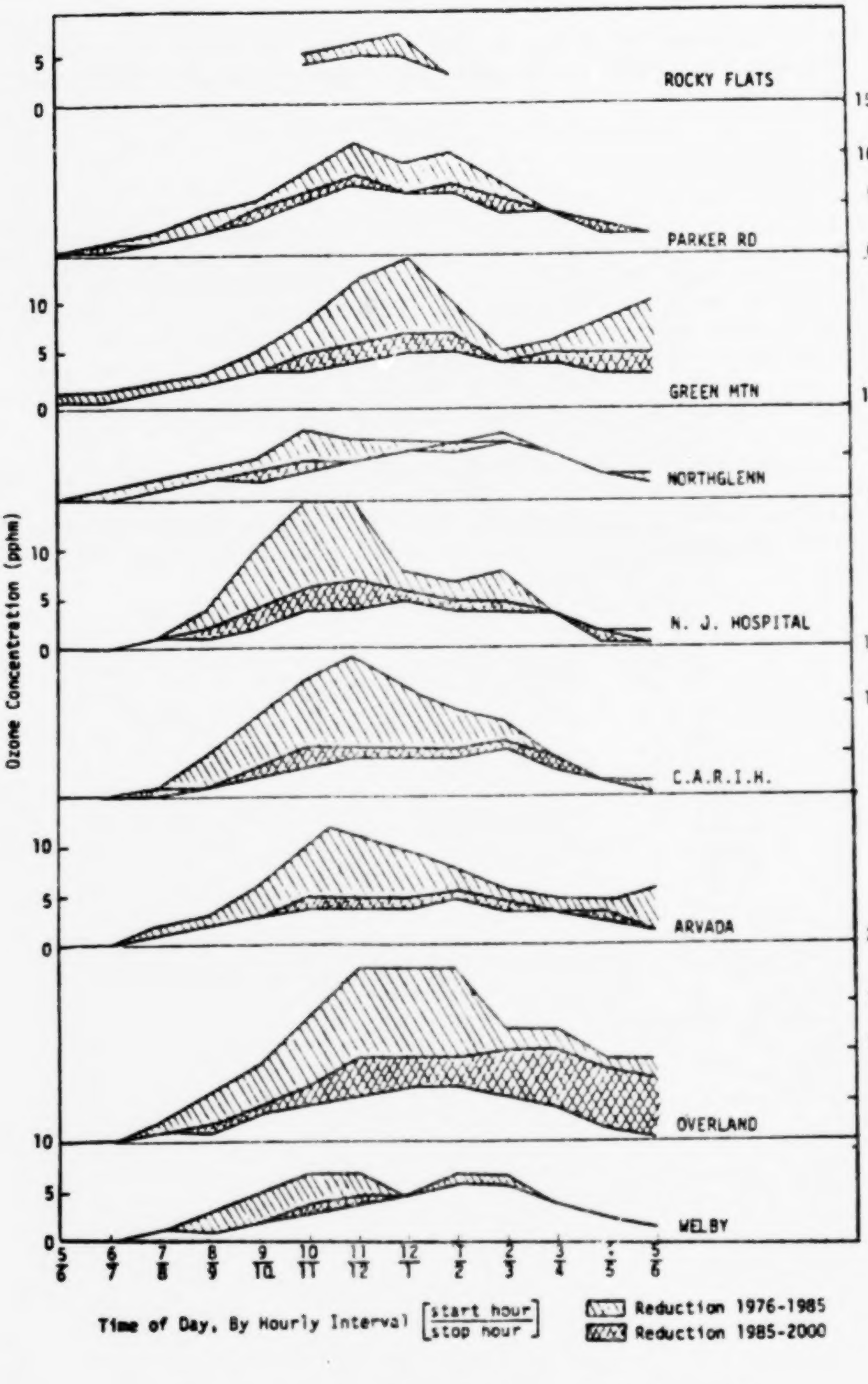
*Does not assume New Source Review Regulation.

[a] These are extrapolated from LIRAQ modeling results.



(a) Meteorology for 28 July 1976 Assumed

Figure 5-6. REDUCTION IN PREDICTED OZONE CONCENTRATIONS (pphm) AT DENVER STATIONS DUE TO PREDICTED FUTURE EMISSIONS CHANGES



(b) Meteorology for 3 August 1976 Assumed

Figure 5-6 (concluded)

- A 65 percent increase in NOx emissions resulted in up to 60 percent reduction in the peak ozone concentration within the modeling grid.

SUMMARY OBSERVATIONS

Both the LIRAQ application in the San Francisco Bay Area and the SAI model application in Denver fell somewhat short of the guidance being offered for model applications. In San Francisco, only one prototype day was used throughout all model applications while in Denver the strategy and emission sensitivity tests were severely limited. Efforts are being made in both regions to improve their respective analyses. In part, the guidance presented in this section is derived from a retrospective evaluation of these two model applications.

EPA is sponsoring additional sensitivity studies for both the SAI model and LIRAQ. (The results of additional LIRAQ sensitivity tests are documented in Volume III of this report.) The guidance offered here should be modified as additional experience in model applications is gained.

REFERENCES

Anderson, G.E., et al., "Air Quality in the Denver Metropolitan Region 1974-2000", Systems Applications, Inc., prepared for U.S. Environmental Protection Agency, Region VIII, EPA-908/1-77-002, May 1977.

Association of Bay Area Governments, Bay Area Air Quality Management District, Metropolitan Transportation Commission, "1979 Bay Area Air Quality Plan", Berkeley, California, January 1979.

Wada, R. Y. et al., "A Methodology for Analyzing Alternative Oxidant Control Strategies", Journal of the Air Pollution Control Association, Vol. 29, No. 4, pp 346-351, April 1979.

6. MODEL INTERPRETATION

Both Lagrangian trajectory and Eulerian grid models are designed to replicate the physical conditions of the atmosphere and the processes which affect pollutant concentrations. This approach differs substantially from previously applied rollback techniques, and leads to practical problems when attempting to interpret model output with respect to the ozone standard. The key problems are summarized as follows:

- The "prototype meteorological days" used in the model performance evaluation may not be the "worst case" days which have occurred, particularly if special field studies are relied upon to collect supplementary data for use in the preparation of model inputs. Since the plan must demonstrate attainment of the ozone standard, which in turn is now statistically defined, some method for relating the effectiveness of control strategies simulated on the prototype days to their effectiveness in meeting the ozone standard must be developed.
- Since no model can be expected to precisely replicate conditions which occurred on a specific day, it may be expected that the model's performance will not be perfect. Therefore, demonstrating ozone levels at or below the ozone standard on the simulated prototype days does not necessarily demonstrate that attainment would occur on the actual days.
- The specification of initial and boundary conditions for simulations of future air quality is a potential problem for cases in which substantial changes in emission distributions and/or levels from the validation case are being tested.

THE "WORST CASE" ISSUE

There are three basic options for resolving the "worst case" issue. The first option is to use a "worst-case" prototype day to evaluate the model's performance in the first place. This, however, is frequently not possible due to a lack of adequate meteorological and air quality data for such a day. The second option is to ignore the issue and presume that the prototype conditions being modeled are reasonable representations of worst case conditions. Methods for extrapolating from prototype conditions to worst case conditions are not well developed at this time, and may introduce additional uncertainties into the modeling analysis. If substantial changes in the magnitude and distribution of precursor emissions are being simulated, it may be argued that the conditions that led to "worst-case" ozone levels on a particular day may not have the same effect under such altered emission conditions. In fact, it is extremely difficult to define in advance what meteorological conditions would lead to "worst-case" ozone levels under emission conditions that are substantially altered from historical

Preceding page blank

patterns. In short, the second option is to avoid introducing additional sources of uncertainty into the modeling analysis.

The third option is to develop a relationship between the prototype days being modeled and the estimated second-worst day. One straightforward method is illustrated in Figure 6-1. Here, the ozone data for the year from which the prototype day(s) was selected is plotted as a standard Larsen-type log-normal distribution. The positions of the estimated second worst day as well as the prototype day may be found on this distribution. Using the same slope of the distribution, the model-predicted ozone level for a given simulation may be used to plot the future year distribution under the emissions conditions being tested. The new distribution can then be used to estimate ozone levels on the second-worst day for that simulation. Here it must be assumed that the distribution of days remains intact in the future. Under this assumption, the somewhat elaborate procedure just described reduces to taking the ratio of the prototype day ozone levels to second-worst case ozone levels, and presuming that the ratio will remain constant in all cases simulated.

This option provides a simple method for extrapolating model output to "worst-case conditions." However, there are a number of conceptual problems with this method. First, by focusing on the maximum level as the measure of air quality on a given day, a great deal of the information being generated by the model is wasted. Second, since the maxima in the future will likely be displaced in both space and time from their original positions, it is likely that the position of a given day on a statistical distribution would also change. Third, portions of the region which may experience small increases in ozone levels as a result of a given strategy cannot be accounted for. Other statistical extrapolation methods could conceivably be developed in the future to overcome these problems.

It may be seen that there are problems associated with each of the options for interpreting model output vis-a-vis the ozone standard. However, a regulatory determination must be made concerning the acceptability of the plan, and its compliance with the requirements of the body of existing laws and regulations. This interpretation of model results is a cornerstone to that determination.

IMPERFECT MODEL PERFORMANCE

There are two basic options for dealing with imperfect model performance. The first option is to accept the model results as they are, making no attempt to "correct" or "calibrate" the results according to ambient monitoring data. If more than one prototype day is being used, and any systematic bias in model results is eliminated during the model performance evaluation phase, then no after-the-fact calibration or correction is necessary. On the other hand, it may not always be possible to eliminate all systematic bias or to obtain satisfactory results for several different prototype days, particularly if schedules are constrained by regulatory deadlines. In such cases, some method

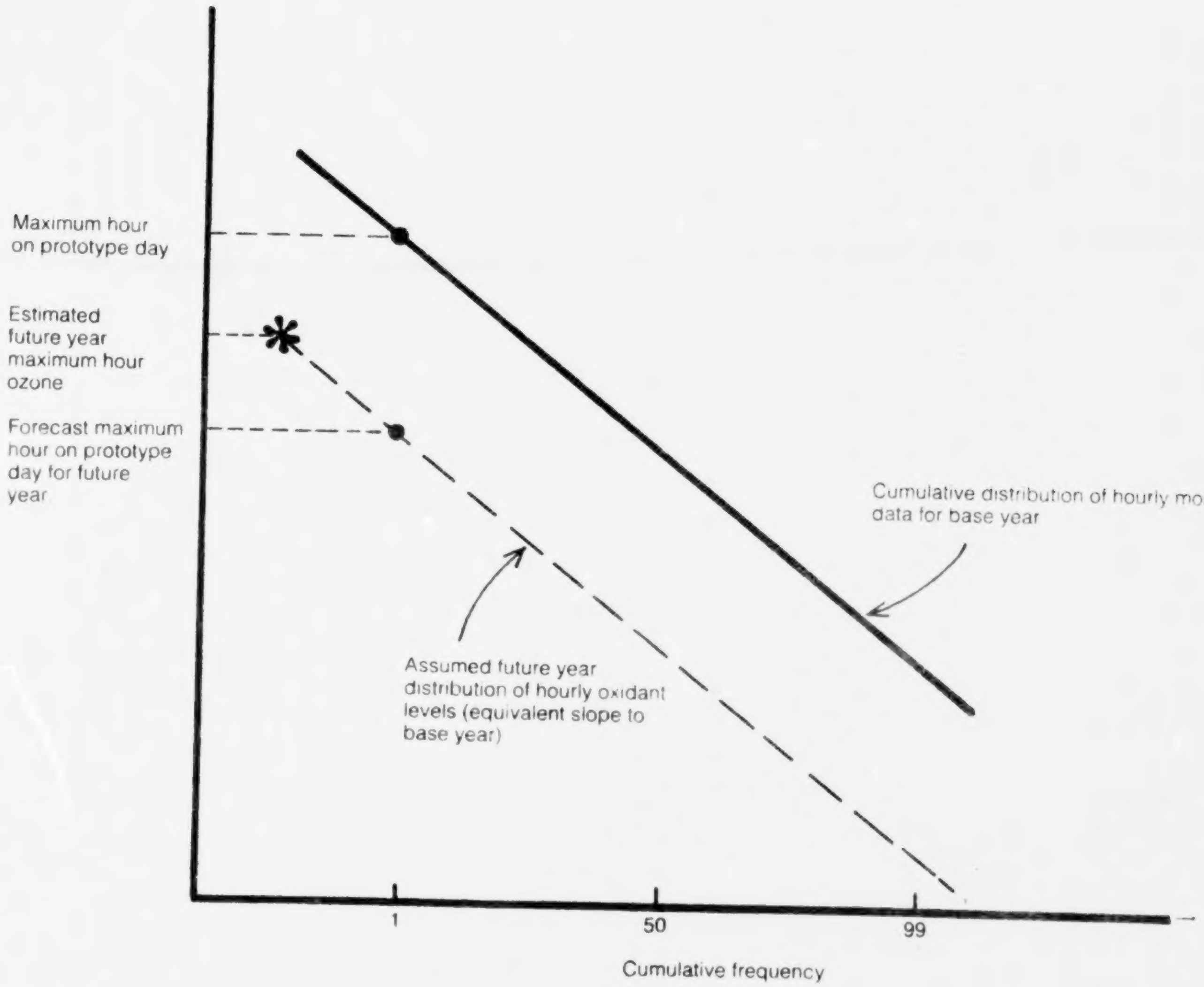


Figure 6-1

Sample statistical extrapolation to determine "worst case" ozone in a future year

should be considered for appropriately compensating the results to ensure that control requirements are not either understated or overstated. For example, a calibration or correction equation may be used as was done in the San Francisco Bay Area example described later.

SPECIFICATION OF INITIAL AND BOUNDARY CONDITIONS

The specification of initial and boundary conditions for simulations of future year conditions is an important problem for ozone modeling, particularly when testing control strategy cases that would reduce the simulated emissions and resulting ozone levels to levels at or near the ozone standard. As emission levels are reduced, the contributions of pollutant concentrations specified at the beginning of a model run (initial conditions) and concentrations specified at the boundaries of the model grid (boundary conditions), become increasingly important to the resulting predicted ozone levels. As previously noted in Section 3, concentrations of pollutants at the boundaries of a metropolitan area are usually poorly known quantities in the modeling analysis--there are very limited data that could be used for model validation on an historical day, and virtually no data for future simulations that could act as a guide. The following examples provide some hints for how to proceed. However, it should be clear that additional research is needed in this area, and that model users should make these assumptions carefully and ensure their full documentation.

METHODS FOR MODEL INTERPRETATION USED IN THE SAN FRANCISCO BAY AREA

The LIRAQ application in the San Francisco Bay Area included the use of an extrapolation to account for worst-case conditions. The existing prototype days in the LIRAQ library were limited by the days for which extensive supplemental monitoring data were available. Those days were not the worst-case days according to historical monitoring data. The maximum one-hour oxidant level of 0.18 ppm recorded on the day for which validation results were best (July 26, 1973), was in the top ten for that year as well as the upper 3 1/2 percent of the overall five year distribution from 1970 to 1974. The expected maximum for that year, however, was 0.24 ppm.

The procedure used to obtain a worst-case evaluation involved a straightforward application of the Larsen model to relate the prototype day to the worst days recorded for the year. Using the daily regionwide high hour oxidant measurement to characterize each day, the distribution of days for the year was developed in the standard Larsen format. By knowing the position of the days according to the monitoring data, it was possible to perform an extrapolation of LIRAQ results to determine an expected maximum oxidant level on the "worse-case day." The procedure involved the use of the maximum one-hour oxidant concentration calculated anywhere on the LIRAQ grid to characterize projected air quality for that day. The resulting adjustment factor was 0.24/0.18 = 1.33.

An adjustment factor was also developed to compensate LIRAQ results for an imperfect validation. Although it was amply demonstrated that LIRAQ exhibited no systematic biases on two different validation days, as shown in Figure 6-2, resource and schedule constraints limited the control strategy analysis to a single prototype day. This left some uncertainty regarding the effect of the control strategies under other meterological conditions.

The formula for the adjustment factor was as follows:

$$\frac{C_f}{C_v} = \frac{C_s}{C_m} \text{ or } C_s = \frac{C_m}{C_v} \cdot C_f$$

where Cm= regionwide high hour ozone concentration measured on the validation day

Cv= regionwide high hour ozone concentration reproduced by the model on the validation day

Cf= regionwide high hour ozone concentration forecasted by the model under some future emission scenario

Cs= regionwide high hour ozone concentration to be computed and compared to the standard.

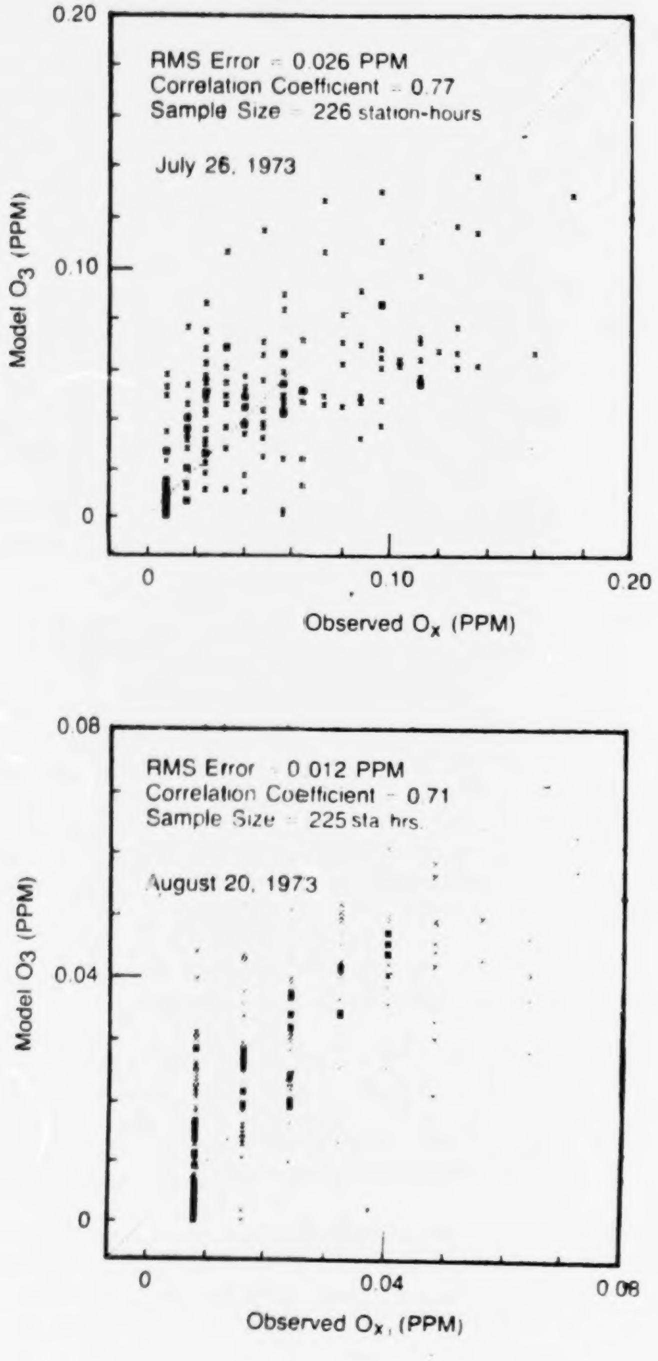
In other words, the ratio of the measured regionwide high hour ozone concentration on a given validation day to the model-produced regionwide high hour oxidant concentration was used to adjust forecasted ozone maxima. This compensated the forecast for any inherent biases in the model or input data. On July 26, 1973 the measured oxidant maximum was .18 ppm, while the model-produced maximum was .17 ppm. The adjustment ratio was therefore .18/.17 = 1.06.

Finally, the problem of specifying initial and boundary conditions for future year simulations was addressed as follows:

- Initial conditions for hydrocarbons and nitric oxide were factored up or down proportionally to the change in the aggregate regional emission inventory for each pollutant. In addition, all simulations were initiated in pre-dawn hours to minimize the initial concentrations of secondary pollutants.
- Since the prototype meteorology being used consisted of a prevailing onshore wind, flow through the upwind boundary was assumed to contain background levels of all species both during model validation and during future year simulations.

Figure 6-2

LIRAQ verification for two 1973 prototype days using 1975 emissions, based on 16 hourly values at 15 locations.



- The vertical boundary condition at the base of the temperature inversion was originally defined to depend partially on the concentration of pollutants in the grid cell below. Therefore, as emission levels changed, the boundary condition at the ceiling would change in the same direction. Since the degree of vertical transport down from the inversion was relatively small on the prototype day, no change in this form of specification was considered warranted.

METHODS FOR MODEL INTERPRETATION USED IN THE DENVER METROPOLITAN AREA

The application of the SAI model in Denver was also constrained by a limited number of days for which supplemental monitoring data were available. Statistical analyses of model performance on the three prototype days indicated no apparent systematic bias in the model results at high ozone levels (see Figure 4-4). Since two prototype days were used in subsequent applications, no correction for an imperfect validation was used.

To estimate second-worst case ozone levels in 1982 and 1987, a correction factor of 1.21 was developed based on the ratio of the actual second highest hour ozone level and the high hour measured on the prototype day. As in the previous SAI study, no correction for imperfect model performance was made. Initial conditions were adjusted for future year simulations by applying the ratio of total regional emissions of each pollutant to the initial early morning ambient concentration field. Since boundary concentrations were assumed to be close to natural background values in all simulations, no adjustment of boundary conditions was made for future year simulations.

The baseline result for 1987, including the worst case correction, was slightly in excess of the .12 ppm ozone standard. For subsequent evaluation of control strategies, linear rollback was applied to the 1987 emissions inventory and model-predicted ozone level to provide a demonstration of attainment. For the 1982 SIP revision, further model applications are planned to evaluate the efficacy of the total control strategy.

CONCLUDING REMARKS

This report was prepared with the expectation that advanced photochemical models will be increasingly used in the development of State Implementation Plan revisions. The guidance and examples presented are based on the limited experience currently available. There are three general problem areas that can be anticipated by potential model users:

- Modeling resources and expertise are diffuse. While there are at least three different modeling groups within EPA, and expertise can be found in various consulting firms, universities, and certain state agencies, there is no single, authoritative and unbiased source for obtaining expert assistance in implementing a model applications program. The potential model user must be able to sort through the literature, obtain advice from a number of different sources, and then decide how to proceed. Moreover, each modeling program will deal with a unique set of conditions, issues, and constraints which will require the exercise of independent judgment and coordination with all participating entities.

o There are no standards for model performance. While attempts are being made to develop model performance standards, a crucial unresolved issue is what to do in the event that a model doesn't perform up to the prescribed standard? All procedures for estimating emission control requirements to meet an ambient standard developed to date involve a model of the relationship between emissions and air quality. One model might perform better than another in a given situation, but it is often difficult to anticipate that in advance, and a substantial investment of time and resources would be necessary to find that out. Until substantial additional experience is gained in model applications under a variety of conditions, or until some reasonable alternative procedure to modeling is developed as a default option, this issue will remain unresolved.

o The validity of model interpretation methods is uncertain. Worst-case corrections, corrections for imperfect model performance, and assumptions concerning future year initial and boundary conditions are complex technical and legal issues in need of further research. Resource constraints may force the use of such interpretation methods in many instances. In each case, the validity of the method should be evaluated in the context of the total model applications program, and its relationship to the plan.

Despite the absence of substantive guidance in each of these areas, the use of photochemical dispersion models will provide the firmest possible technical foundation for SIP development, and will represent a substantial improvement over previously used methods. As additional experience is gained in model applications, the guidance that can be offered should become more specific and useful to both modelers and model users.

APPENDICES

A0

APPENDIX A

PREPARATION OF METEOROLOGICAL INPUT FIELDS FOR LIRAQ SIMULATIONS IN THE SAN FRANCISCO BAY AREA

The LIRAQ model has been used in several studies of air quality in the San Francisco Bay Area. Because this area is characterized by complex and changing flow and mixing-height fields, the preparation of suitable meteorological inputs to the model is difficult and time-consuming. This Appendix illustrates some of the factors that had to be considered in order to prepare the necessary meteorological input fields.

Selection of days was originally largely determined by availability of data from extensive field studies. Even though the permanent meteorological observation network in the Bay Area is quite extensive, the mesoscale circulations are sufficiently complex that the field studies were judged to be essential in order to characterize the wind and mixing height fields in the region. Later, after experience had been gained, it was possible to develop prototype days for which there were no supplemental field measurements.

Model Grid Regions. The 170 x 120 km region that can be studied with the LIRAQ model is shown in Figure A-1. The MASCON meteorological preprocessor code (see below) is limited to a maximum field of grid elements of 65 by 65. The whole Bay Area can be treated with a 5-km grid. For smaller (1 x 1 km or 2 x 2 km) grid elements, the region must be subdivided into several subregions. The photochemical version of the model (LIRAQ-2) is limited to a 20 x 20 grid (e.g., a 100 x 100 km area with 5 km grid elements). The choice of a particular 100 x 100 km subregion depends upon a) the area whose air quality is to be studied, b) the locations of the source emission regions that affect the study area, c) the minimization of boundary effects, and d) possible flow reversals whereby pollutants might be transported out of the region and then returned later.

Meteorological Data Sources. Upper wind and temperature data were available from the National Weather Service (NWS) Oakland upper-air station (twice daily) and from San Jose State University. Surface data were gathered from 85 permanent sites on designated field observation days. Most of these sites provided surface wind and temperature data. The surface observation sites are operated by many organizations, including the Bay Area Air Quality Management District (BAAQMD), various airports, the California Division of Forestry, various military bases and private industry. Solar radiation measurements were obtained from 9 BAAQMD monitoring stations. In addition to all of these "quasi-permanent" data sources a large body of supplemental information was collected from special field studies for some of the prototype days that were developed. Vertical profiles of temperature and several pollutants, obtained from aircraft spirals at several Bay Area locations, were very useful in specifying the time- and space-dependent inversion base height field and upper boundary conditions.

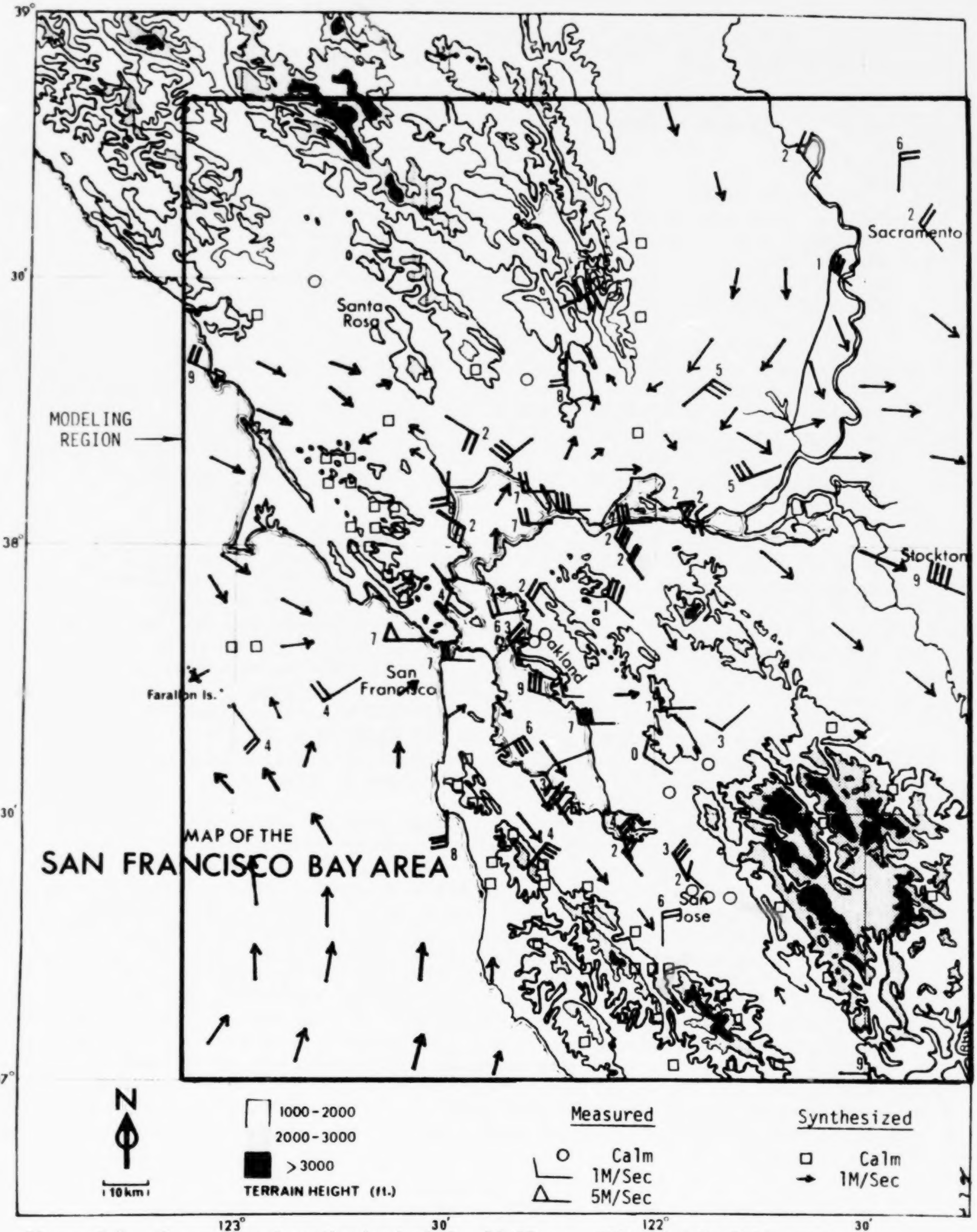

39°
30′
38°
30′
37°
123°
30′
122°
30′
MODELING REGION
MAP OF THE
SAN FRANCISCO BAY AREA
Santa Rosa
Sacramento
Stockton
Oakland
San Francisco
San Jose
Faralon Is.
N
10 km
1000-2000
2000-3000
>3000
TERRAIN HEIGHT (ft.)
Measured
Calm
1M/Sec
5M/Sec
Synthesized
Calm
1M/Sec

Figure A-1. Observed and synthesized winds, 10:00 a.m. PST, 24 July 1974, used as inputs to MASCON.

The Mass-Consistent Wind Field Model (MASCON). The MASCON computer code (Dickerson, 1978) generates mass conserving fields of wind velocity and of inversion base height from observed and analyst-synthesized input data. The code interpolates irregularly spaced input data to a regular grid and employs an adjustment scheme based on variational analysis to the mass fluxes (mean winds within the layer times mixing depths) to ensure agreement with an appropriate form of the continuity equation.

MASCON was developed because visual interpretation of the meteorological fields is difficult, especially for the Bay Area's complex terrain and changing meteorology, and therefore semi-objective procedures were needed to simplify the analysis. Results are not always consistent with expectations; consequently application of the mass-consistent data is usually an iterative process. When a particular analysis reveals areas with unrealistic divergence or flow patterns, a reanalysis with additional or adjusted synthesized wind or inversion data must be performed before proceeding to simulate air quality.

When a constant density mixing layer is assumed, mass is not allowed to accumulate or to deplete within a given zone (defined as a single grid square extending vertically from the surface to the spatially and temporally varying inversion height). This constraint must be satisfied for the transport calculations, which are based on the flux into and out of a given zone. If this constraint is not satisfied, the mass gained or lost within a zone as a result of an imbalance of the net fluxes can invalidate any calculated transport of diffusion of air pollution concentrations.

Because LIRAQ is developed in a flux form, the appropriate equation of continuity can be written as (Dickerson, 1978):

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} + \frac{\partial (hv)}{\partial y} + w = 0,$$

where H is the height of the inversion base above topography, u and v are the east-west and north-south components of the mean velocity within the mixed zone, respectively, and w is a vertical velocity. When observational data (or data interpolated from observations) are used for h, u, v, and w in a finite difference form of the above equation, a non-zero residual (ε) appears on the right-hand side. In some cases small values of ε can produce errors in mass fluxes which, for LIRAQ, are large enough essentially to invalidate the transport, eddy diffusion, and chemical transformation calculations. Therefore it is necessary to ensure that the ε field over the grid is sufficiently small that the resultant fictitious mass losses or mass gains are acceptably small.

MASCON is designed to adjust the product of the mean wind and the height of the inversion base above topography in such a way that the air mass is conserved and the observational data are changed as little as possible. The adjusted wind and inversion fields, prepared at three-hour increments, are then used as input meteorological data for LIRAQ.

Wind Field Inputs to MASCON. Observational data tend to be concentrated in the populated areas and to become much more sparce in the rural regions and at the outer edges of the domain of interest, particularly over the ocean. Because the model needs wind data at each grid point, it utilizes a Gaussian weighting (according to distance R of the observation from the grid square) interpolation scheme in order to develop a regional wind field from the relatively sparce observed data field. In some regions such interpolation can lead to unrepresentative winds at some grid points far from any observed wind vector. Thus, wind vectors at some remote locations are best synthesized based on subjective meteorological experience and in consideration of such factors as described below.

One of the first steps in synthesizing wind data is a determination of the gradient wind over the Bay Area and an assessment of how this should be reflected in the subinversion layer. In the absence of other information this gradient wind may be used as an estimate of the wind around the boundary of the model area. Variations of this estimate over the model area should be made only while one is simultaneously considering the effects of topography, the land and sea breeze regimes, and mass conservation in the subinversion layer.

In order to prepare a reasonable pattern of observed and synthesized winds such as that shown in Figure A-1,the analyst must be aware of the nature of the mesoscale circulations in the planetary boundary layer over the San Francisco Bay Area. These flows are influenced by coastal hills and gaps and by the thermal contrasts that are set up by differential heating over the ocean, bays and adjacent land areas. On most summer and early fall days there is an elevated subsidence inversion with a layer of marine air beneath. In the early morning,thermal gradients are weak and weak drainage flows prevail out of the coastal valleys. In the afternoon the marine layer advances onshore, driven by a sea breeze circulation that results from the pressure difference between cool coastal waters and the hot interior valleys. For a fully developed sea breeze condition there is pronounced channeling of the winds through the coastal gaps and into the interior valleys. Past statistical studies of local climatological flow patterns (Smalley, 1957 and MacCracken, 1975) were found to be an invaluable aid in evaluating mesoscale flow characteristics for preparation of synthesized winds.

Inversion Base Height Inputs to MASCON. The topography of the inversion base height is one of the most important data fields needed by the LIRAQ model because it influences such factors as mixing depth, wind profiles, pollutant concentration profiles, and the height to which topography controls the flow. Factors to be considered and used as guidelines in constructing regional inversion-base-height contours (for use as input to MASCON) from limited amounts of observed data are described here.

Over the adjacent part of the Pacific Ocean, surface temperature is essentially uniform and thus also is the depth of the marine mixed layer (and therefore the height of the subsidence inversion). Inland, temperature gradually departs from temperatures over the ocean as the trajectory distance increases from the ocean. This difference, due to solar heating and radiational cooling of the land surface, undergoes a diurnal cycle with the land temperature exceeding the ocean temperature in late afternoon.

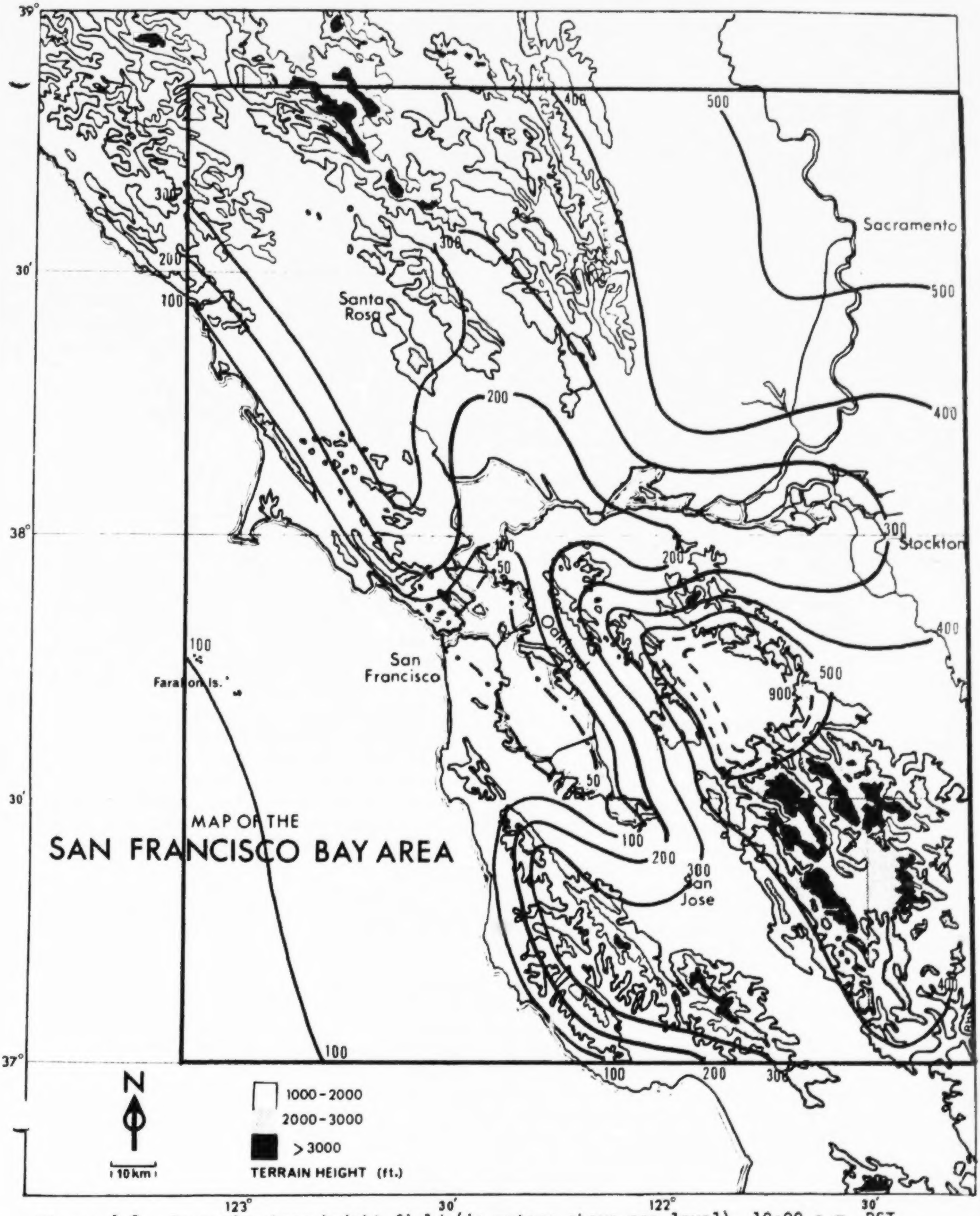
At this time, if it is not completely destroyed, the inversion will reach its maximum height. Conversely, the land will cool below the ocean temperature at night, and a new low-level radiation inversion will form that will reinforce the subsidence inversion.

Because of radiation effects, elevated surfaces act as heat sources during the day and heat sinks at night, thus enhancing the diurnal temperature cycle and its effect on the inversion base height. For modest topographic features the inversion-base-height contours should tend to parallel topography contours unless overpowered by the effect of trajectory distance from the ocean. Furthermore, when the onshore flow becomes moderate-to-strong, the air tends to retain the temperature and inversion-base-height characteristics of the oceanic source region. Thus, in the absence of other effects, inversion-base-height contours will tend to parallel air trajectories. In general, in regions where there are strong horizontal gradients in the inversion-base-height or flow fields, the streamlines and inversion-base-height contours will tend to be parallel.

Before putting sparse wind and inversion base height data into MASCON, it is necessary to have experienced meteorologist construct fields of wind velocity and of inversion base height (at three-hour intervals throughout the period to be simulated) in order to allow additional synthesized data to be provided to the interpolation scheme. Examples of these fields are shown in Figures A-1 and A-2. Because of this need for competent human intervention in a region as complex topographically as the Bay Area, the initial processing of wind and inversion base height fields is quite time-consuming, requiring several person-weeks of effort by an experienced meteorologist. In a less complex region, this task could be greatly simplified.

Once the wind and inversion base height fields, at three-hour intervals, have been prepared and put into MASCON, an interative process begins. First, MASCON interpolates the wind and inversion base height fields for each grid cell averaged over three-hour periods. Graphical output is also produced, as shown in Figures A-3 and A-4. Next, MASCON adjusts the fields until the flow fields are mass conserving. Figure A-5 shows the adjusted (mass-consistent) flow field derived from the inversion-height and wind fields shown in Figures A-3 and A-4, respectively. Finally, the meteorologist studies the adjusted flow field and if it appears unrealistic in any way, he modifies the input fields and runs MASCON again. The procedure may be repeated several times before acceptable mass-consistent flow fields are obtained.

Preparation of Fields of Atmospheric Transmissivity. The photochemical module of LIRAQ is very sensitive to the variation of the photodissociation rates. These rates are proportional to a wavelength-dependent photon flux density. The "clear sky" photon flux density at the earth's surface was calculated as a function of zenith angle. This value was then multiplied by a transmission coefficient representing the ratio of Eppley pyranometer measurements to calculated clear sky transmission. Thus the transmission coefficient is a measure of solar flux obstruction by clouds and aerosol.



39°
30'
38°
30'
37°
123°
30'
122°
30'
Sacramento
Santa Rosa
San Francisco
Oakland
San Jose
Stockton
Faralon Is.
MAP OF THE
SAN FRANCISCO BAY AREA
500
400
300
200
100
50
900
N
10 km
1000-2000
2000-3000
> 3000
TERRAIN HEIGHT (ft.)

Figure A-2. Inversion base height field (in meters above sea level), 10:00 a.m. PST, 24 July 1974, constructed as a guide for generating input values to MASCON.

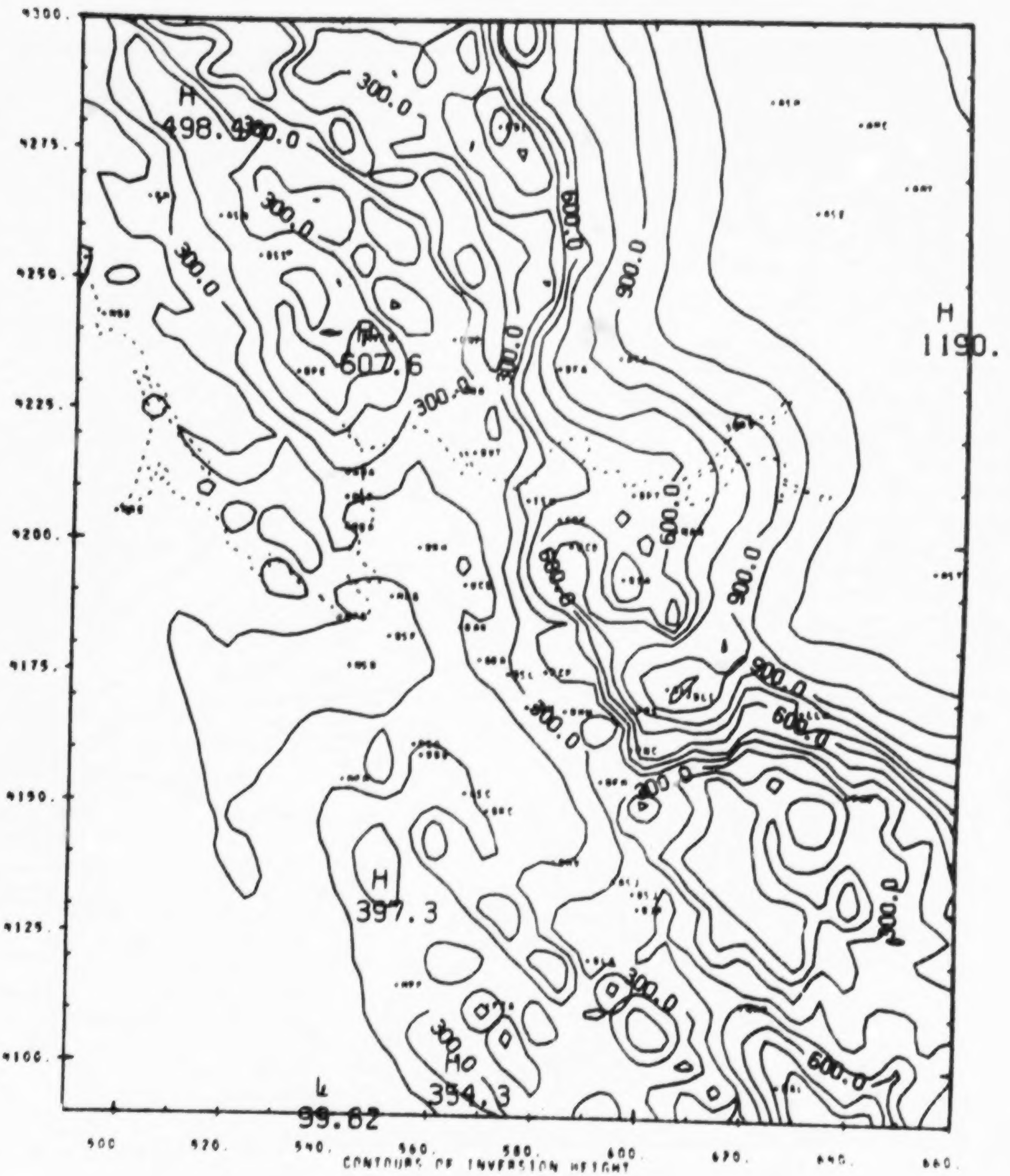

4300.
4275.
4250.
4225.
4200.
4175.
4150.
4125.
4100.
500
520.
540.
560.
580.
600.
620.
640.
660.
H
498.4
300.0
H
607.6
H
1190.
600.0
900.0
H
397.3
H
354.3
L
99.62
CONTOURS OF INVERSION HEIGHT
MEASURED ABOVE TOPOGRAPHY

Figure A-3. MASCON-generated interpolated field of mean inversion base height above topography (meters), averaged over three hours (10:00 to 13:00 PST), 24 July 1974.

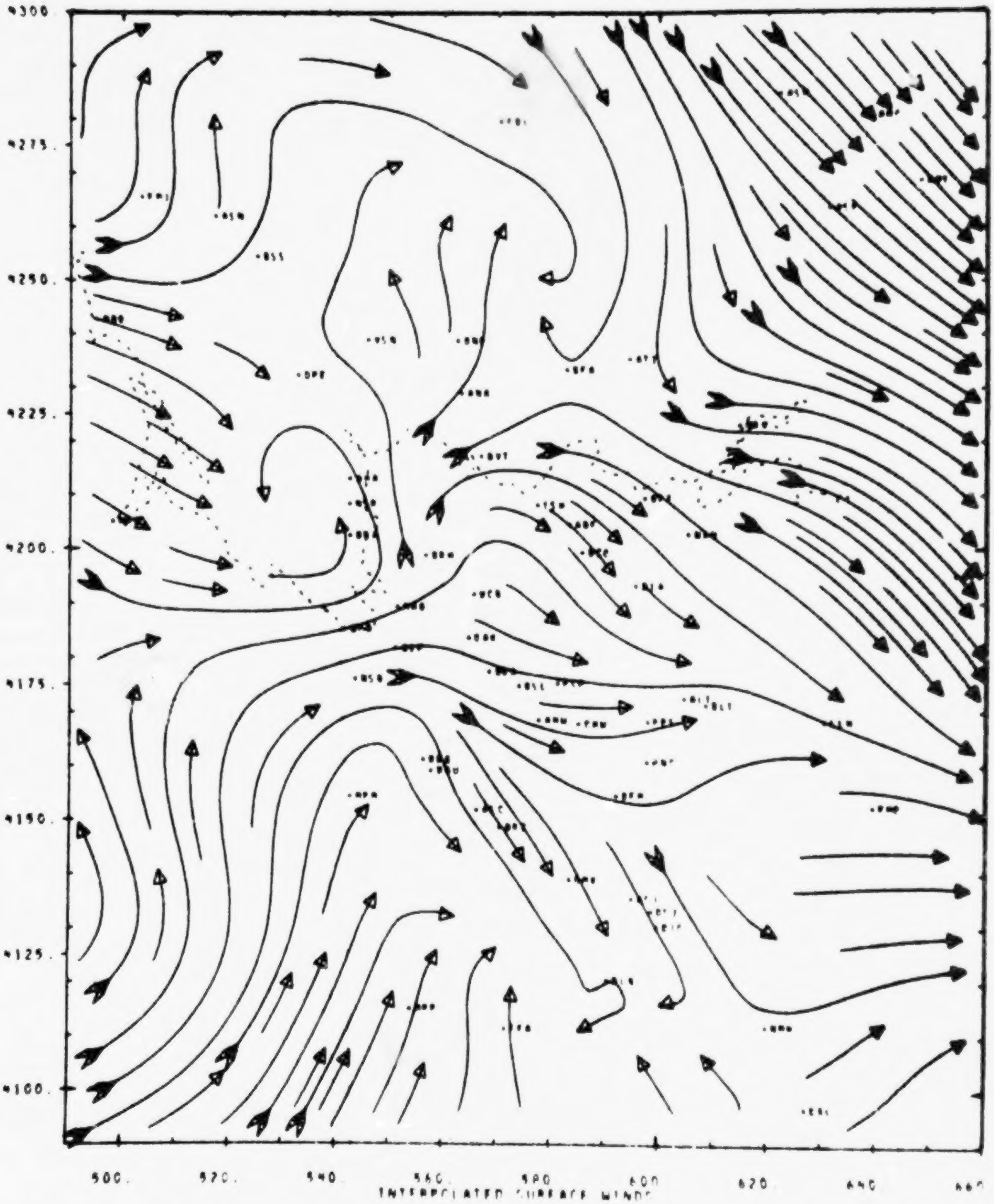


Figure A-4. MASCON-generated interpolated field of surface winds, averaged over three hours (10:00 to 13:00 PST), 24 July 1974.

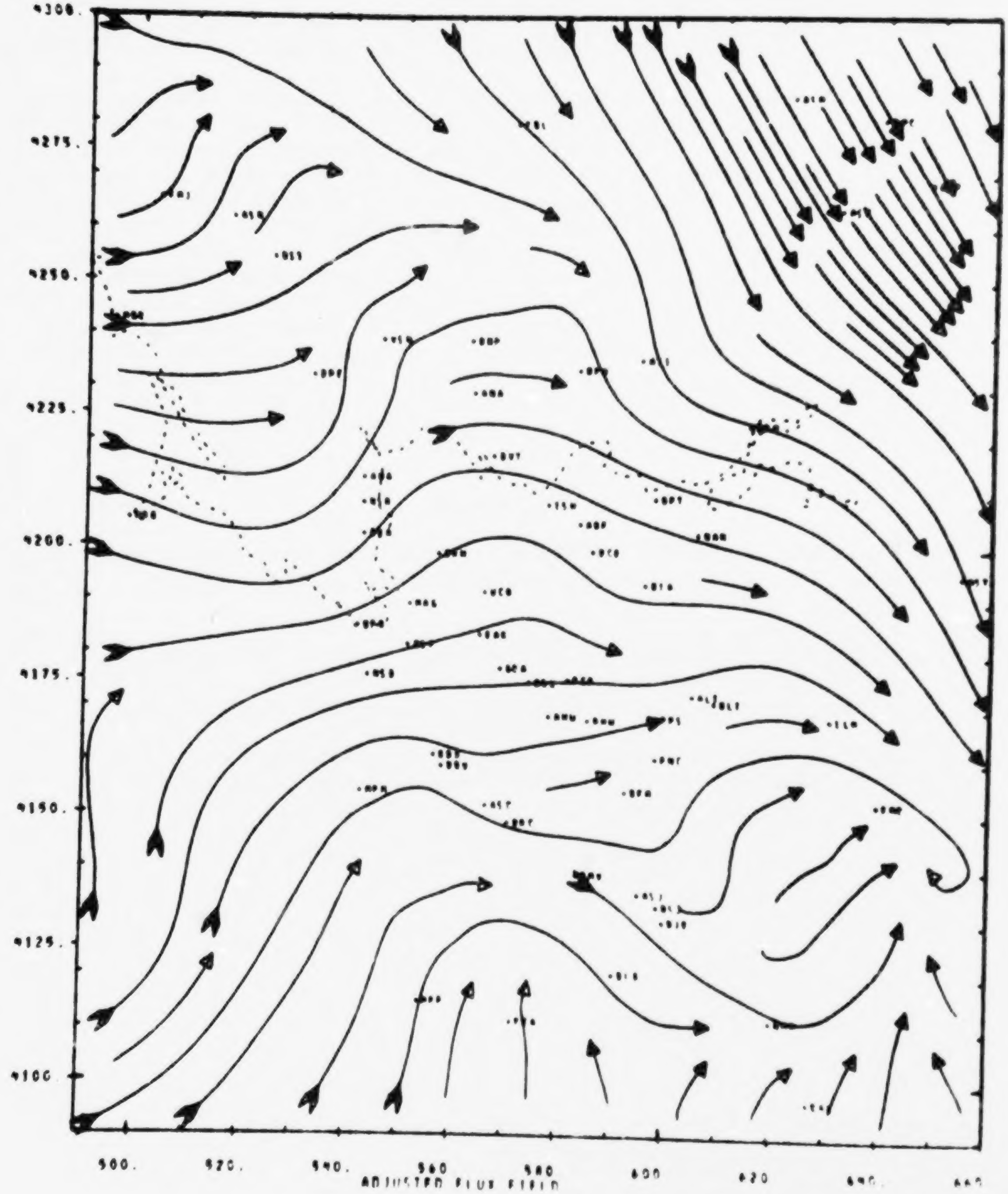


Figure A-5. Mass-consistent flow field derived from the interpolated inversion base height and wind fields shown in Figures A-3 and A-4.

The transmission coefficient fields were estimated from pyranometer measurements made at nine San Francisco Bay Area stations operated by the BAAQMD. These transmission coefficients were averaged for the three-hour intervals over which the LIRAQ model assumed uniform wind fields.

Analysis over the mountainous regions of the Bay Area was a problem due to the absence of data there. Because the transmission coefficient probably increased with elevation, it was decided that the mountainous regions above 2000 ft and covering a significant area would be bordered by a contour two to four units higher than the surrounding lower elevation region. The analysis of the August 20, 1973 maps was further complicated by the presence of stratus just offshore and over portions of the Bay. This stratus failed to appear over any of the pyranometer stations, and consequently no estimates of transmission coefficients typical of such cloud cover were available. To handle the situation transmission coefficients representative of stratus overhead for each 3-hour period were determined by comparison with pyranometer data from San Francisco on a day with stratus. These were then used in the August 20 analyses. A sample plot of the regional fields so developed is presented in Figure A-6.

References

1. Dickerson, M.H. (1978), "MASCON - A mass consistent atmospheric flux model for regions with complex topography," J. Appl. Meteor., Vol. 17, No. 3, pp 241-253.

2. MacCracken, M.C., and G.D. Sauter, eds. (1975), "Development of an Air Pollution Model for the San Francisco Bay Area, " UCRL-51920, Vol. 2, Lawrence Livermore Laboratory, Livermore, California.

3. Smalley, C.L. (1957), "A Survey of Airflow Patterns in the San Francisco Bay Area," Preliminary Report, United State Weather Bureau Forecast Center, San Francisco, California.

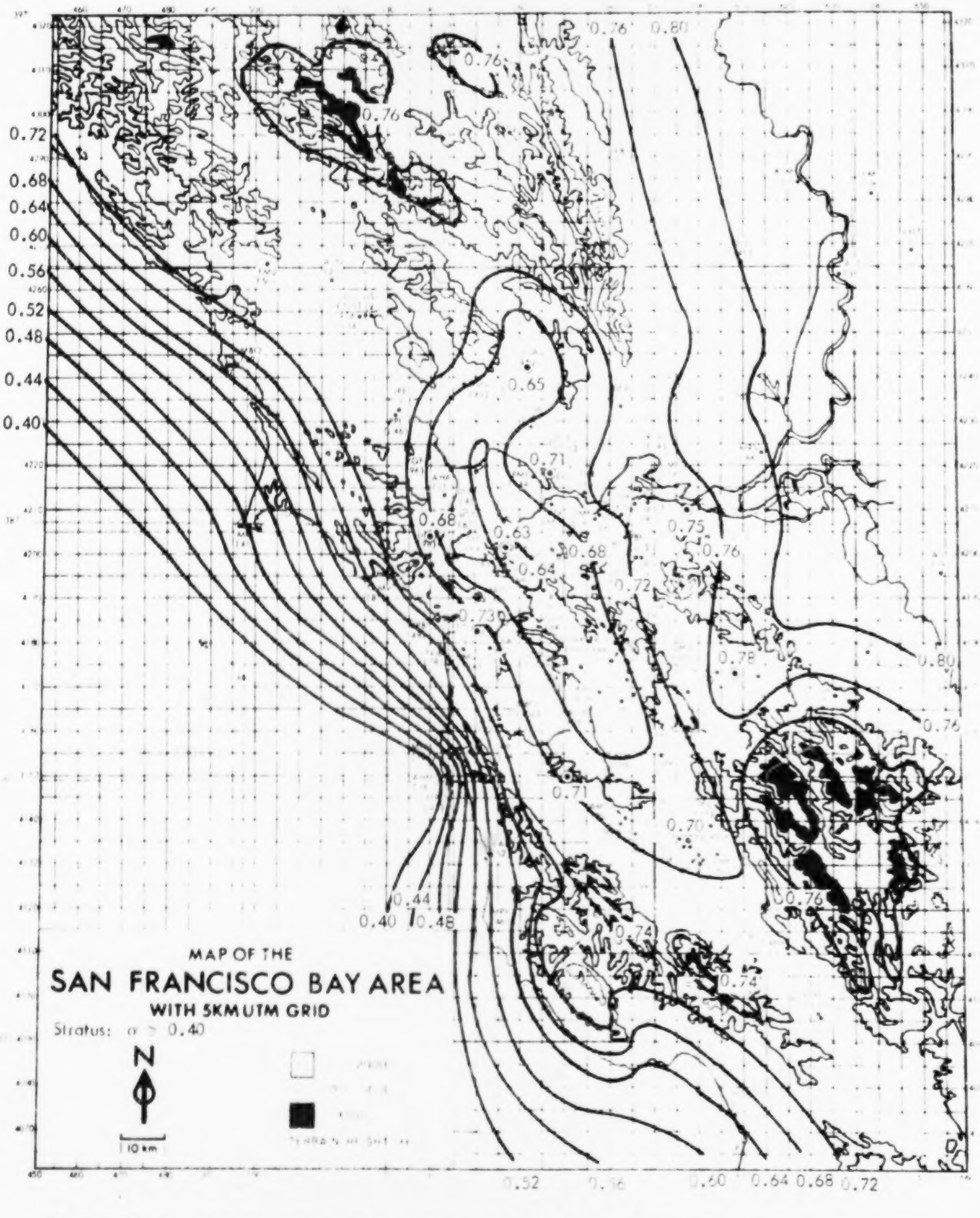

MAP OF THE
SAN FRANCISCO BAY AREA
WITH 5KM UTM GRID
Stratus: α ≥ 0.40
N
10 km
0.72
0.68
0.64
0.60
0.56
0.52
0.48
0.44
0.40
0.76
0.80
0.76
0.76
0.65
0.71
0.68
0.63
0.68
0.64
0.72
0.75
0.76
0.73
0.78
0.80
0.76
0.71
0.70
0.76
0.44
0.40
0.48
0.74
0.74
0.52
0.56
0.60
0.64
0.68
0.72

Figure A-6. An example of the transmission coefficient fields prepared for MASCON input. This analysis represents conditions observed on 20 August 1973, 13:00 to 16:00 PST. Source: MacCracken and Sauter, 1975.

Reproduced from best available copy.

APPENDIX B

PREPARATION OF METEOROLOGICAL INPUT FIELDS FOR SAI SIMULATIONS IN DENVER AND LOS ANGELES

The SAI Airshed Model has been used in several air quality studies in the Denver and Los Angeles areas. Each of these regions has its own distinct topographical, meteorological and air quality characteristics. This Appendix illustrates some of the problems encountered by the modelers and describes how these problems were resolved. In each region, as experience with the model increased, it was possible to refine some of the methodologies to produce more realistic meteorological input fields and thereby improve model performance.

1. DENVER SIMULATIONS

Selection of days to simulate was based on episode days on which special field studies had been mounted to collect additional air quality and meteorological data. On one of these days a frontal system moved through the region, causing air pollution levels to be lower than expected. The maximum ozone concentration observed during the summer of 1976 was 27 pphm; however, supplemental meteorological data were not collected on this day. Instead, three days including a peak ozone level of 17 pphm were selected because supplemental field measurements were available for these days.

The modeling grid in the original Denver modeling work reported by Donnely (1978) was a 30x30 mile portion of the Denver metropolitan area, using 1x1 mile gridded emissions. Later, Anderson, et al. (1977) found that simulations using 2x2 mile grid squares produced nearly identical results. Anderson, et al. also noted that pollutants were sometimes advected out of the modeling region and then returned later when the wind reversed. To contain the Denver pollutant "cloud" within the model, later simulations carried out by Reynolds et al. (1979) used a 60x60 mile area. However, few meteorological observations were available in the outlying areas of the expanded grid.

Meteorological data sources included upper-level winds and mixing depths derived from twice-daily radiosondes from the National Weather Service (NWS) upper air station at Stapleton International Airport. The NWS also provided hourly measurements of surface winds, surface temperature, humidity, visibility and cloud cover. Hourly surface winds were obtained from 24 sites operated by the Colorado Division of Highways (CDH), the Air Pollution Control Division of the Colorado Department of Health, the NWS and various private organizations. Total solar radiation data (hourly averages) were collected at one CDH site. Hourly surface temperatures were obtained from 5 monitoring locations.

Wind field inputs to the Airshed Model were estimated from hourly surface observations. Data from up to 24 surface monitoring sites were used to construct hourly estimates of the wind speed and direction for each ground level grid cell. These wind fields were generated by the inverse-distance-weighted interpolation scheme described by Liu, et al. (1973). Because there were relatively few wind observations in some of the outlying areas of the grid, it would otherwise be necessary to extrapolate using the measured values in the interior of the grid region. To maintain some control over this extrapolation process, additional "fictitious" stations were defined and assigned wind speeds and directions based on an analysis of the actual observations and the local terrain features. The resulting surface wind fields were smoothed by replacing the wind velocity in each grid cell with the vector average of the wind velocities in the block of nine cells (in the Denver city area) or 25 cells (in outlying areas) centered at the cell of interest.

In the original Denver simulations, winds aloft were set equal to the surface wind field. In more recent simulations, a procedure developed by Killus, et al. (1977) was employed which produces a wind profile consisting of a mean velocity that is a function only of height and a component resulting from local convergence or divergence effects.

Mixing depths were assumed to vary only as a function of time because there were insufficient data with which to characterize spatial variations in the mixing depth field. Hourly values of the mixing depth were established using the two temperature soundings made at Stapleton International Airport and available surface temperature observations. During the predawn hours on both simulation days, the early morning (5:00 a.m.) temperature sounding indicated the existence of both a surface and an elevated inversion layer. Based on an examination of the temperature records, it was estimated that the surface based inversion was completely eroded by 7:00 a.m. and 8:00 a.m. on 29 July 1975 and 28 July 1976, respectively. After these hours, mixing was confined to the height of the elevated inversion base. To estimate this height, at each hour during the day, the surface temperature aloft was extrapolated at the vertical gradient exhibited in the afternoon temperature sounding. This methodology predicted that the elevated inversion layer was broken up by about noon.

Other meteorological variables included hourly relative humidity and ambient temperature data obtained from the NWS station at Stapleton International Airport. These values were used to derive estimates of the water vapor concentration. Vertical temperature gradients, both below and within the elevated inversion layer, were derived from the available radiosonde observations. The temperature at the inversion base was estimated by extrapolating the surface temperature aloft at the gradient indicated in the afternoon temperature sounding. Hourly values of the exposure class (used to estimate vertical diffusivities) were derived from estimates of the diurnal variation of solar insolation. The values of photolysis rates were increased by 15 percent to account for Denver's elevation of over 5000 ft, above sea level. Finally, the atmospheric pressure input was set at 0.85 atm.

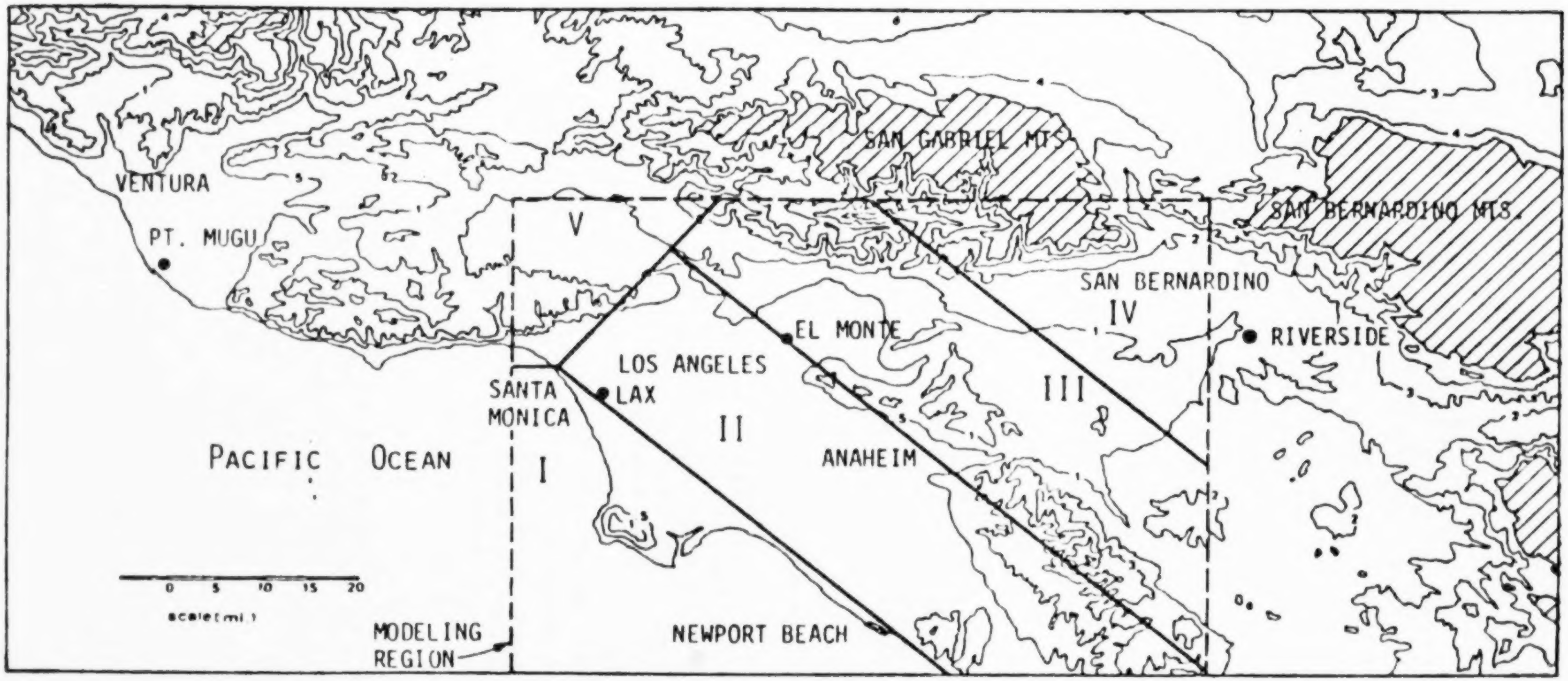
2. LOS ANGELES SIMULATIONS

Prototype day selection. Two ozone episodes that occurred in Los Angeles during the summers of 1974 and 1975 provide the basis for carrying out the Airshed Model simulations for this large metropolitan area. During the 26-29 June 1974 episode, a land-sea breeze regime prevailed--a common occurence in the South Coast Air Basin (Kieth and Selik, 1977). At night, pollutants were advected out over the water; with the ensuing morning sea breeze, some of this material was reintroduced into the air basin. The 26 June meteorological conditions resulted in the occurence of high (34 pphm), but not extreme, peak ozone levels. These conditions are representative of a "typical" smoggy day in Los Angeles. A simulation was also made of 27 June, when the highest ozone concentration (49 pphm) observed during 1974 occurred.

On 4 August 1975, during another high ozone episode, onshore winds prevailed throughout the entire 24-hour period. In fact, analysis of the synoptic meteorology and local wind station measurements revealed persistent onshore winds for the two days preceding 4 August as well. Selection of this day allowed evaluation of model performance for a characteristically different set of meteorological conditions.

An important consequence of the unusual meteorological regime on 4 August 1975 is the greater ease of preparation of the three-dimensional wind fields and initial and boundary conditions for that day compared with the 1974 episode days (Tesche and Burton, 1978). To facilitate the specification of boundary conditions for the 1974 episode days, it was necessary to include as much of the Pacific Ocean as possible in the modeling region to enable the treatment of offshore pollutant transport. For the 4 August simulation day, boundary conditions reflect essentially "clean" marine air, unperturbed by local anthropogenic emissions sources. In this case it was not necessary to include grid cells over the Pacific Ocean in the modeling region, since there was no significant amount of offshore transport.

Selecting the modeling region. Before discussing the procedures used to generate the various model inputs, we first consider the definition of the modeling region. At the time the simulations reported here were carried out, the computer programs that composed the Airshed Model would only accommodate a region subdivided into about 900 grid squares. Gridded emissions data were available for an 80x40 array of 2x2 mile grid squares covering the area shown in Figure B-1. As a result, it was necessary to model only a portion of the large 160x80 mile region. After considering the locations of the highest measured ozone concentrations and the need to include portions of the Pacific Ocean to treat offshore pollutant transport, a 36x25 array of 2x2 mile grid squares was selected, encompassing the area indicated in Figure B-1.


VENTURA
PT. MUGU
SAN GABRIEL MTS.
SAN BERNARDINO MTS.
SAN BERNARDINO
IV
RIVERSIDE
V
EL MONTE
LOS ANGELES
LAX
SANTA MONICA
II
III
PACIFIC OCEAN
I
ANAHEIM
0 5 10 15 20
scale(mi.)
MODELING REGION
NEWPORT BEACH

SOURCE: Reynolds, et al., 1979

FIGURE B-1 INTERPOLATION REGIONS FOR THE PREPARATION OF MIXING DEPTHS

Meteorological data sources included upper level winds at Riverside, El Monte, Los Angeles International Airport (LAX), Pt. Mugu and San Nicholas Island. Mixing depths were determined from radiosondes from LAX, El Monte, Pt. Mugu and San Nicholas Island, from an aircraft spiral at Riverside and an acoustic sounder at El Monte. Hourly surface winds were available from up to 60 stations operated by the county air pollution control districts, the California Air Resources Board (CARB), the NWS and the Pacific Missile Test Center (PMTC) at Pt. Mugu. Hourly surface temperatures and humidity were available from 15 stations. In 1974, hourly radiation measurements were made at LAX, downtown Los Angeles and Riverside. In 1975, radiation was measured at El Monte and Riverside. Visibility and cloud cover data were available from seven airports.

Wind fields. To ascertain how model performance is influenced by the procedure employed to generate wind inputs, Reynolds et al. (1979) prepared three different wind fields for the 26 June 1974 simulation day. Starting with the surface and aloft wind measurements, three different methods were used to calculate a three-dimensional wind field. Two of these methods, (1) interpolation using the station values, and (2) use of a three-dimensional wind model produced acceptable results.

o Interpolation Technique. Preparation of the three-dimensional wind inputs required by the Airshed Model involved several steps. Wind data from 60 locations in the Los Angeles Air Basin were obtained from the California Department of Transportation, the CARB, and the local air pollution control districts. The 34 stations within the modeling region and the stations just outside of the region were used to construct a surface wind field using an inverse distance weighted interpolation scheme as described by Liu, et al. (1973). The resultant wind field was smoothed to eliminate any sharp changes in wind speed and direction. The smoothing algorithm replaces each grid value by the five point average about that grid cell.

Winds aloft were estimated using the three wind soundings made in the morning--at El Monte, Riverside, and LAX--and two made in the afternoon--at El Monte and LAX. These measurements taken aloft were used along with the ground-level measurements to estimate an upper level "synoptic" wind velocity for each hour. This wind velocity was assumed to determine completely the wind field above the mixing layer.

For grid cells situated between the surface and the top of the mixed layer, the winds were determined by interpolation, using the information from the surface field and the synoptic wind. In making estimates, it was assumed that the divergence in the wind is surface-generated and that this divergence dies out with height, being completely dissipated above the mixing layer.

o Three-Dimensional Wind Model. The three-dimensional wind model used in this study is described by Yocke, Liu, and McElroy (1977) and Tesche and Yocke (1978). The model computes horizontal flow fields at up to 10 equally spaced heights, starting at sea level. These heights are measured from sea level as opposed to ground level; thus, if a region contains mountains at any particular level, no wind vectors will be calculated in those areas. The flow fields for each level are calculated by specifying the surface temperature field; surface roughness; average ground elevation for each grid square; and the winds around the boundary of the region, estimated from the available wind data.

The wind field obtained from this model is mass consistent and highly influenced by topography. The general direction of the flow for a particular level is largely determined by the terrain configuration, surface heating patterns, and general features of the synoptic level flows. In this study it was difficult to establish the wind flows aloft for each hour because instantaneous upper air soundings were made at a maximum of six stations only once or twice a day.

In the Airshed Model, the z-direction has been transformed so that all heights are measured from the ground rather than from sea level. This means that the output from the wind model must be transformed so that the wind vectors are specified at the nodes of the Airshed Model grid cells rather than at the equally spaced heights above sea level used by the wind model. Consequently, some of the features of the original wind field are lost; most importantly, the transposed horizontal wind field is not mass consistent. A suitable vertical wind velocity is calculated within the Airshed Model at the time the pollutant calculations are carried out which, in effect, renders the wind field mass consistent. For Los Angeles the final wind field of this three-dimensional model reflects the channeling of the wind through the mountain passes more realistically than does the interpolated wind field model.

Mixing Depths. Special consideration must be given to the preparation of mixing depths in Los Angeles because the behavior of the inversion over that area is complicated by the differing heating and cooling rates of the air over the land compared with that over the ocean. To account for this complexity, the modeling region was divided into subregions, as shown in Figure B-1. The subregions were drawn parallel to the coastline in conformity with the notion that inversion height is reasonably constant along lines parallel to the Los Angeles coastline (Edinger, 1959). The data that were available to estimate mixing depths were morning and afternoon soundings at LAX and El Monte, a morning sounding at Riverside, hourly acoustic soundings at El Monte, three daily soundings at Pt. Mugu, and hourly surface temperatures at these five locations. These observations were used to estimate hourly mixing heights for LAX, El Monte, and Riverside.

The LAX values were used for Subregion I, the Riverside values were used for Subregion IV, and the grid squares along the border of subregions II and III were given the El Monte values. The mixing depths for Subregions II and III were then interpolated using the two closest assigned values. The mixing depths for Subregion V were also interpolated, using the closest subregion on the east and the Pt. Mugu value on the west.

After every grid square had been assigned a mixing depth according to the above scheme, the mixing depth field was modified in the vicinity of significant terrain features. The scaling factor is represented by the ratio of the mixing depth to the sum of the ground elevation and the mixing depth. This factor is designed to have a pronounced effect in reducing the mixing depth at high ground elevations, but a relatively small influence as the mixing depth decreases. The values of the mixing depths at measurement sites were altered so that, upon application of the scaling factors, the measured mixing depths were obtained.

Once the mixing depth has been determined, the height of the modeling region can be calculated. The modeling region consists of six layers of grid cells, with the first layer being the "surface layer" with a constant height of 60 feet. At any point in the region, the other five layers are equal to each other in height. For the time period 5:00 to 8:00 a.m. PST, four of these five layers were within the mixed layer and one was above; for the time period 8:00 a.m. PST to the end of the simulation, all five of the layers were within the mixed layer.

Other Meteorological Variables. Temperature and humidity were measured at five stations within the modeling region. The observations for each hour were averaged, so that a single temperature and a single water vapor concentration were used as inputs to the model. Temperature gradients, both above and below the inversion base, were used to characterize atmospheric stability and to estimate plume rise from large point sources. These inputs were established from the available temperature soundings.

References

1. Anderson, G.E., et al. (1977), "Air Quality in the Denver Metropolitan Region: 1974-2000," EPA-908/1-77-002, Systems Applications, Incorporated, San Rafael, California.

2. Donnelly, D.E. (1978), "Oxidant Model Applications: Denver," 57th Annual Transportation Research Board Meeting, January 1978, Washington, D.C.

3. Edinger, James G., (1959), "Changes in the Depth of the Marine Layer Over the Los Angeles Basin," *J. Meteorol.*, Vol. 16, No. 3, pp. 219-226.

4. Kieth, R.W., and B. Selik (1977), "California South Coast Air Basin Hourly Wind Flow Patterns," South Coast Air Quality Management District, El Monte, California.

5. Killus, J.P., et al. (1977), "Continued Research in Mesoscale Air Pollution Simulation Modeling - - Vol. V: Refinements in Numerical Analysis, Transport, Chemistry, and Pollutant Removal," EF77-142, Systems Applications, Incorporated, San Rafael, California.

6. Liu, M.K., et al. (1973, "Further Development and Evaluation of a Simulation Model for Estimating Ground Level Concentrations of Photochemical Pollutants - - Vol. III: Automation of Meteorological and Air Quality Data for the SAI Urban Airshed Model, " Systems Applications, Incorporated, San Rafael, California.

7. Reynolds, S.D., et al. (1979), "Photochemical Modeling of Transportation Control Strategies, Vol. I. Model Development, Performance Evaluation, and Strategy Assessment," EF79-28, Systems Applications, Incorporated, San Rafael, California.

8. Tesche, T.W., and C.S. Burton (1978), "Simulated Impact of Alternative Emissions Control Strategies on Photochemical Oxidants in Los Angeles," EF78-22R, Systems Applications, Incorporated, San Rafael, California.

9. Tesche, T.W., and M.A. Yocke (1978), "Numerical Modeling of Wind Fields over Mountainous Regions in California," Conf. on Sierra Nevada Meteorology, American Meteorological Society, 19-21 June 1978, South Lake Tahoe, California.

10. Yocke, M.A., M.K. Liu, and J.L. McElroy (1977), "The Development of a Three-Dimensional Wind Model for Complex Terrain," *Proc. of the Conference on the Applications of Air Pollution Meteorology*, 28 November - 2 December 1977, Salt Lake City, Utah.

APPENDIX C

Preparation of Emission Inventory For Photochemical Modeling In the San Francisco and Denver Regions

The LIRAQ source inventory for the San Francisco Bay Area is made up from four component parts, each part compiled with independent data sources and techniques. The four components as shown in Figure C-1 are: major point sources, area sources, airports, and mobile sources. Major point sources include oil refineries, electric utilities, chemical industry, metallurgy, rock and mineral operations, etc. -- any stationary source emitting more than 0.1 ton/day or 25 tons/year of any pollutant. Such sources are listed separately in the existing source inventory, with information on location, emissions, stack parameters, operating schedules, and process variability. Area sources, also called "population-distributed" emissions, include: domestic fuel combustion, off-road mobile sources, utility engines, and small stationary sources such as service stations, dry cleaners, small plastic manufacturing, etc. Emissions are estimated by a variety of techniques including direct measurement, natural gas use, solvent sales, gasoline sales, point and resin use, etc. Airports include emissions from commercial, military and general aviation aircraft from 37 airports in the Bay Area. Table C-1 presents the relative contributions of the four source components for the 1975 baseline inventory. It is clear that mobile and area sources are the largest organic emission contributors. These two categories are also the most complex for emissions estimates and spatial and temporal resolution. Spatial and temporal distribution techniques used in the Bay Area study are discussed below, for each component of the inventory.

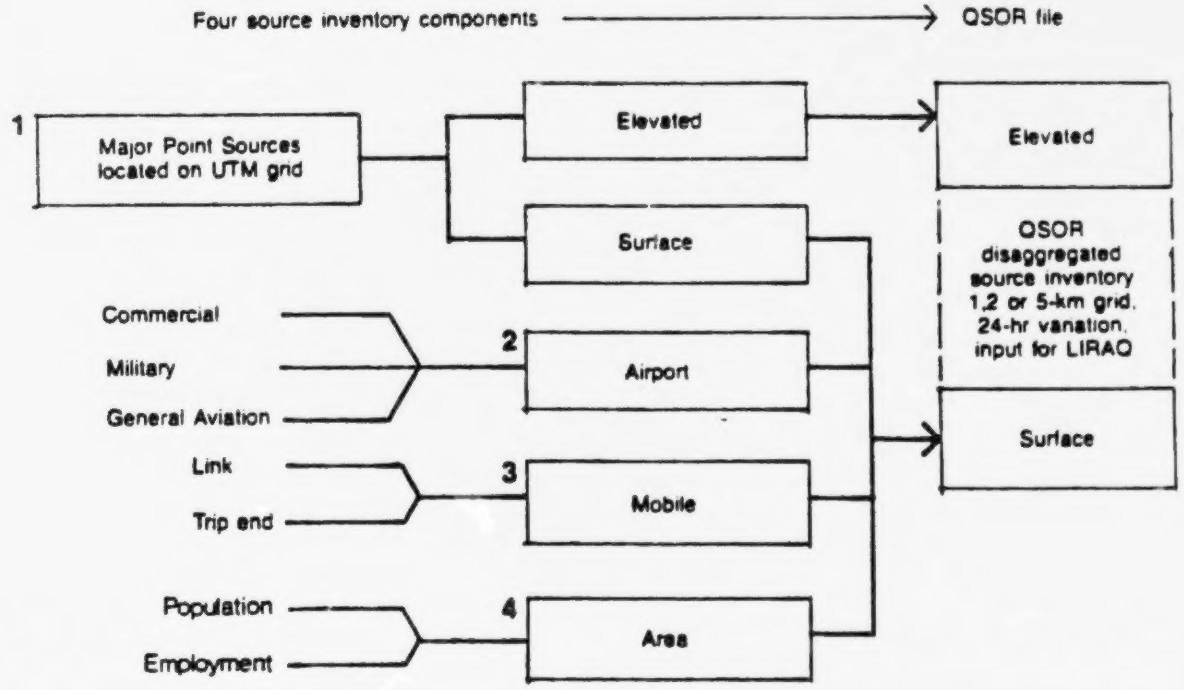
Major Point Sources

Data concerning emissions from more than 120 major industrial operations in the Bay Area are maintained as part of BAAQMD's source inventory. Data maintained include: UTM coordinates, stack height, actual measurements of stack gases, operating hour, seasonal variations, and fuel usage records. Based on these records, the total emissions from major point sources were distributed into appropriate UTM grid cells. Major point sources are carried as separate listings and are further divided into "surface" and "elevated" categories, depending on the stack height. (The dividing line is 100 feet.) The elevated major point sources remain as individual listings in the final source inventory file, QSOR. Surface major point sources are merged with emissions from the other three source inventory components.

The hourly and seasonal variations of major point sources are generally known. Because these are large individual emitters, they are subject to intense scrutiny from regulatory agencies. And by the nature of large operations, they usually maintain reliable internal process records. Based on operating hours and seasonal variations, hourly emission distributions for each major point source in the Bay Area were constructed.

Figure C-1

Summary schematic of QSOR[a] file preparation.


Four source inventory components
QSOR file
1
Major Point Sources located on UTM grid
Elevated
Surface
Commercial
Military
General Aviation
2
Airport
Link
Trip end
3
Mobile
Population
Employment
4
Area
Elevated
QSOR disaggregated source inventory 1,2 or 5-km grid. 24-hr variation. input for LIRAQ
Surface

---

[a] QSOR is the code name for the final disaggregated source inventory file which serves as input to the Livermore Air Quality Model.

TABLE C-1

1975 Source Inventory for San Francisco Bay Region by Source Categories

Source Type / Source Category	Part.	Org.	NO_x	SO_2	CO
		(tons/day)			
Major Point Source	19	119	167	187	74
Area Source	98	412	101	12	338
Airport Source	9	20	13	1	55
Stationary Total	126	551	281	200	467
Mobile Source	43	472	399	19	3,869
Total	169	1,023	680	219	4,336

Source: BAAQMD

Airports

The locations of 37 airports in the Bay Area were identified by UTM coordinates. Because some landing and take-off emissions are spread out within the mixing layer, airport emissions were distributed over neighboring grid squares. For commercial and military flights, emissions were distributed uniformly over all grid squares within 2 miles of the airport. For general aviation at community airports, a distance of 1 mile was used.

Aircraft emissions are divided into three source inventory classifications: commercial carriers, military, and general aviation. Commercial carriers are most important from an emission viewpoint, and these are also the best documented. Comprehensive schedule books provide up-to-date listings of commercial flights to all the major airports of the world, with arrival and departure times and aircraft types. For the LIRAQ project, data were compiled on commercial carrier operations at Bay Area airports from the 1975 "North American Air Guide." Temporal resolution factors for three commercial airports (San Francisco, Oakland and San Jose) were based on these data.

Operations data for four military air bases in the Bay Area were very limited and also difficult to predict. Emission estimates were based on fuel usage data, but actual daily flight schedules were not available. Based on available data, hourly operations for four military air bases were estimated to be 90% during daylight hours and 10% during night flights.

Finally, general aviation emissions were uniformly distributed over daylight hours for 28 small community airports, and extended to include some early morning and late evening flights at busier airports. Hourly traffic counts on airport approach roads were used to check the diurnal pattern of general aviation.

Area Sources

Among the 107 activity classifications in the BAAQMD source inventory, 58 include some area source contributions. Emissions in 1975 from area sources comprised about 58% of total particulates, 39% of organics, 15% of nitrogen oxides, 6% of sulfur dioxide, and 8% of carbon monoxide emitted in the Bay Region.

The method developed for spatial resolution of area sources in the Bay Region was a "cross classification" technique. The objective was to postulate functional relationships between source categories and a variety of demographic, economic, and land use variables being used as surrogates. A table of coefficients was compiled to link 58 source activity classifications (those with area source components) with 19 known employment categories from ABAG's "Series 3 Projections." (See Hoffman, et al., 1978, and Perardi, et al., 1979).

The Series 3 projections cover population, housing, employment and land use in the Bay Area. For the nine counties around San Francisco Bay, the data are compiled for 440 sub-regional areas termed "zones," which are made up of one to approximately seven 1970 census tracts. Housing is recorded by dwelling unit, and population/employment by 23 categories. The information was based on census data, local surveys, fertility and immigration statistics. A list of the Series 3 variables used in this project is provided in Table C-2. Since the Series 3 data were internally consistent between the 1975 base year and future year projections, use of this data ensured that the area source emissions projections and spatial distribution would also be consistent.

Before they could be used as a basis for are source distribution, the Series 3 data had to be distributed over the 1-km UTM grid system. This critical step was accomplished by a combination of manual and computer techniques. First, regional maps were used to eliminate those grid squares which are essentially uninhabited. Those areas (bays, tidelands, marshes, mountains, etc.) comprise about 75% of the total area. Series 3 variables were then distributed from 440 zones to the remaining grid squares, which total 5000 to 6000 km2 of developed or developable land. The exact total depends on the year being considered.

A cross-classification table was then developed to link certain types of area sources with appropriate Series 3 variables. For some source classifications a direct correspondence could be found. For example, BAAQMD source category no. 18 "Farming Operations" could be linked with Series 3 employment category P7 "AGRI" which includes agricultural production and services. Similarly, source classification no. 40 "Printing" could be distributed with Series 3 "MFG1" which is printing, publishing and related industries. In most cases, however, the source classification did not fit clearly with a single Series 3 variable. For these cases, professional judgment was employed to produce a multiple distribution formula, so that area source emissions from a single source classification could be distributed with two or more Series 3 variables. For example source classification no. 35 (Degreasers) provides area emissions of 42 tons/day of organics. These were distributed as follows: 60% with MFG5 (fabricated metal products), 20% with RET. SERV. (including auto repairs), 10% with MFG4 (including electrical and optical equipment), and 10% with OTHER SERV. (including local transit and transportation services). Excerpts from the classification table are shown as Table C-3. The percentage values were selected based on knowledge of local industry and operating conditions.

Area source emissions were distributed and then totalled for each Series 3 category (for each pollutant). Totals were divided by the known total population of the category to produce a per capita emission rate. As an example, for the 394 tons/day of organics for area source distribution, the total for Series 3 category P9, from all source classifications, was 20.75 tons/day. The total employment population in P9 (printing and publishing) was 25170, so the per capita emission factor was .00082 tons/day of organics per printing publishing employee. The per capita emission rates, for each Series 3 category and each pollutant, were then used with the known Series 3 population distributions to produce the

Table C-2 Summary of list of nineteen Series 3[a] variables used in cross-classification analysis (for spatial resolution of area source emissions).

Variable Code	Variable Name	SIC[b] Classification	Description
P1	DWELL	(not applicable)	Dwelling units
P7	AGRI	1, 7-9	Agriculture, forestry
P8	MIN	10, 13, 14	Mining, quarry, oil & gas extraction
P9	MFG1	27	Printing, publishing
P10	MFG2	26, 28, 29, 32, 33	Petrol., chem., paper, metal industries
P11	MFG3	20	Food and kindred products
P12	MFG4	19, 36, 38	Electrical, optical, machinery & instr.
P13	MFG5	34, 35, 37	Fabricated metal products
P14	MFG6	22-25, 31, 39	Textiles, apparel, wood, leather
P15	TRAN	40, 42, 44-46	Transportation (non-auto), pipelines
P16	WHOL	50, 52	Wholesale trade, building material
P17	FIN	62, 63, 67	Financial, insurance
P18	SERV 1	73	Business services
P19	SERV 2	82, 84, 89	Educ. service, museums, galleries
P20	GOV	91, 92	Government
P21	RET	53-59	General merchandise & food stores
P22	BUS. SERV.	80, 81, 96	Health, legal, admin. services
P23	RET. SERV.	70, 72, 75-79	Hotels, personal service, repairs
P24	OTHER SERV.	15-17, 41, 47-49, 60,	Construction, transit, utilities, banking,
	" "	61, 66, 93-95, 99	real estate, other

---

[a] ABAG Projections of population, employment, etc.

[b] Standard Industrial Classification Manual 1972.

Table C-3 Excerpts from the cross-classification table used for spatial distribution of area source emissions.

Tabled values represent the percentage of the area source emissions (from a given source classification) to be distributed with the indicated Series 3[b] variable. Blanks are zeros.

Area Source Classification			Series 3 Categories							
No.	Description	tons[c]	1 Dwell. Units	7 Agric. Forest	9 Print. Publish	11 Food Prod.	12 Elec/Opt. equip.	13 Fabr. metal	23 Retail serv.	24 Other serv.
18	Farming operations	-		100						
19	Food/agric. proc.	6.1				100				
29	Org. solv. storage	10.9					20	20	5	
31	Indus.coating, solv.	99.2			10		10	50	5	
35	Degreasers	42.4					10	60	20	10
36	Dry cleaning,perc.	13.9							100	
40	Printing	10.2			100					
87	Lawn	5.5	100							

[a] The full cross-classification table has 58 area source classifications and 19 Series 3 variable categories. (Numbering is not serial.)

[b] ABAG Series 3 Projections of population, employment, etc.

[c] Area source organics emissions, tons/day, for a summer weekday. Other pollutants have different emission rates but use the same distribution percentages.

area source spatial resolution. Results were checked by summing area source emissions over all grid squares and comparing the resulting total with the total area source emissions used as the starting point.

Changes in the percentage values shown in Table C-4 do not change the amount of area source emissions (as long as the entries sum across to 100%). Only the distribution of the emissions would be changed.

The hourly distribution of area source emissions is based on diurnal variation coefficients (percent of total daily emissions per hour of the day) for each source classification. Weighted hourly variation factors were then produced by multiplying area emissions per classification times the diurnal variation factors of each classification. The resulting (normalized) set of factors were then used for temporal resolution of all area sources emissions.

Mobile Source Emissions

In the San Francisco Bay Area, the highway, or "link" related, hot-stabilized emissions were computed using modified versions of the two Federal Highway Administration computer codes SAPOLLUT and SAPLSM, previously mentioned. The trip end related emissions were computed using programs developed at ABAG. The overall sequence of operation and input data requirements and sources for both codes are summarized in Figure C-2. As shown, each set of programs outputs hourly emissions that are geographically distributed by one kilometer UTM grid squares. The two data sets are then merged for input to the air quality model (LIRAQ). For input, both codes require transportation data from the Metropolitan Transportation Commission (MTC) and emission factors from the California Air Resources Board (ARB). A summary of baseline transportation data inputs is shown in Table C-5.

The motor vehicle emission factors used were derived through the use of a California Air Resources Board emission factor program, EMFAC3. This program was, in turn, based on EPA's Supplement 5* to AP-42, with some minor modifications.

The EMFAC3 program computed hot-stabilized emission factors for HC and NOx in units of grams per mile. It provided emission factor estimates for average route speeds from 5 to 50 mph, ambient temperatures from 20 to 80°F, and any desired mix of cold and hot start operation. Factors were produced for both a weighted average of four vehicle types (light duty auto, light duty truck, heavy duty gasoline, heavy duty diesel), and individully for each vehicle type.

---

*This supplement has been replaced by "Mobile Source Emission Factors," EPA-400/9-78-005, March, 1978.

Table C-4

AREA SOURCE DISTRIBUTION PERCENTAGES FOR SERIES 3 ACTIVITY CATEGORIES

Reproduced from best available copy.

Source Classification	Series 3 Categories* (1 – 23, OTHER LOC.)
CHEMICAL	[redacted]
[illegible], Chem. Ind.	
[illegible] Ind./Agr.	[redacted]
Metallurgical	
[illegible]	100
Mineral/Stone, Sand	[illegible]
[illegible]	[redacted]
[illegible]	10, 10
[illegible] Mineral Prod.	15, 50
[illegible] Operations	100
[illegible] Prod.	100
[illegible] Mill	5, 10, 70
Wood Products Mfg.	10, 80, 10
[illegible]	
GASOLINE DISTRIB.	[redacted]
Vehicle Fill Sta. - -Spillage	
- Storage Tanks	5, 3
[illegible]	5
ORG. COMP. EVAP.	[redacted]
Storage Tanks - -Solvent	
-Misc. Org. Comp.	5, 10, 50, 11, 17, 10, 10, 10, 5, 5
.Coat.-Solv.Base	10, 10, 10, 10, 10
.Coat.-Water Base	10
Com./Dom. Coat. -	[redacted]
-Solv. Base	
-Water Base	60, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3
[illegible]	10, 60, 3, 3
Dry Cleaners-PERC	20, 10
Dry Cln.-Misc.Solv.	100
Rubber Fabrication	100, 100
Plastic Fabrication	40
Printing	100, 10
Misc. Org. Evap.	5
COMBUSTION OF FUELS	[redacted]
[illegible] Oil	100
Com.&Inst. - Gas	5, 10, 5, 10, 10, 25, 10, 10
- Oil	
Power Plts.Coal Boil.	100
[illegible] - Nat. Gas	100
[illegible] - Oil	100
[illegible] - Gas	
[illegible] - Oil	20, 10
[illegible] Ind. Comb. -	[redacted]
-Gas	
-Oil	
[illegible] OF MATERIALS	[redacted]
[illegible]	
Com. Inst. [illegible]	
Ind. [illegible]	10, 20, 20, 10, 20, 20, 5, 5, 10, 10, 20, 10, 20, 10, 10
Agric.Debris Burn	100
[illegible]	100
Accident.Wild Fire	100
Accident.Struc.Fire	3
OFF-HWY MOBILE	[redacted]
Agric. Tractors - -Gasoline	
-Diesel	
Construc.Equip. -	[redacted]
-Gasoline	10, 5
-Diesel	50, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 10, 3
Steamships,Cargo/Pas.	TREAT AS POINT SOURCES IN SHIPPING LANES
Steamships,Tankers	
Steamships,Military	
Cargo Motorships,Fuel	
Motor Tankers,Fuel	
Diesel Motorships - -Cargo/tugs	
Diesel Motorships - -Tanker	
Fishing/Ferryboats	
Lt.Craft [illegible]	100
[illegible]	100
Misc. Utility Engs.	20, 10, 20, 10, 10, 10, 10, 10

*Series 3 categories defined in Table C-2.

Source: BAAQMD

TABLE C-5. SUMMARY OF BASELINE TRANSPORTATION DATA INPUTS TO MOTOR VEHICLE EMISSIONS ESTIMATION FOR THE SAN FRANCISCO BAY AREA

PARAMETER	YEAR			
	1965	1975	1985	2000
AVERAGE WEEKDAY VEHICLE TRIPS				
o Homebased work	1,706,983	2,144,693	2,542,951	3,038,406
o Non-work	5,370,480	6,904,098	8,215,373	9,859,449
o LDV Total	7,077,463	9,048,791	10,758,324	12,897,855
AVERAGE WEEKDAY VEHICLE MILES				
o Homebased work	14,055,453	20,199,644	23,645,050	30,309,087
o Non-work	27,873,495	40,623,164	52,516,997	73,350,341
o LDV sub-total	41,928,948	60,822,808	76,162,047	103,659,428
o HDV @ 12.8%	5,366,905	7,785,319	9,748,742	13,268,407
o Total VMT	47,295,853	68,608,127	85,910,789	116,927,835

ARB's EMFAC3 was the basis for computing the mobile source emission factors.** However, a number of variables, which vary with geographical location and situation, can affect emissions estimates considerably: average vehicle speed, ambient temperature, type of vehicle, percentages of cold and hot start trips and percent of travel by vehicle age (see Figure C-2). Therefore, localized correction factors reflecting these variables were developed based on Bay Area conditions.

Speed and ambient temperature correction factors were also developed from formulas provided in EPA's Supplement 5 to AP-42. Estimates of link speeds and the distribution of vehicle types were provided by the Federal Highway Administration. Assumptions for the diurnal variation of average ambient temperatures were estimated from average summer minima and maxima observed at different regional locations. The ambient temperature correction factor was insensitive to temperatures above 80°F. Vehicle age distributions and pollution control equipment deterioration rates (provided by EMFAC3) were also incorporated into the emission factors.

Link Emissions

A coded highway network (for 1975 and updated for 1985 and 2000) and a transportation model to forecast travel volumes on each link were the basis for the link emissions calculations. As previously stated, a modified version of SAPOLLUT (U.S. DOT, 1976) was used to actually compute the link emissions, given the appropriate link information and the emission factors. The modified model computed estimates of speed on the highway network according to the volume/capacity ratio on each link, for each hour. These speed estimates determined the appropriate speed correction factor to apply.

The model also provided diurnal traffic distributions and the distribution of vehicle types on different road types. Five types of vehicles were examined: light duty autos (LDA), light duty trucks (LDT), heavy duty gasoline-powered vehicles (HDG), heavy duty diesel-powered vehicles (HDD) and motorcycles. The distribution of total VMT among the various vehicle types was obtained from ARB as follows:

Vehicle Type	Percent of Light Duty Vehicle VMT (LDA + LDT)
LDA	86.2%
LDT	13.8%
HDG	8.6%
HDD	4.2%
Motorcycle	0.9%

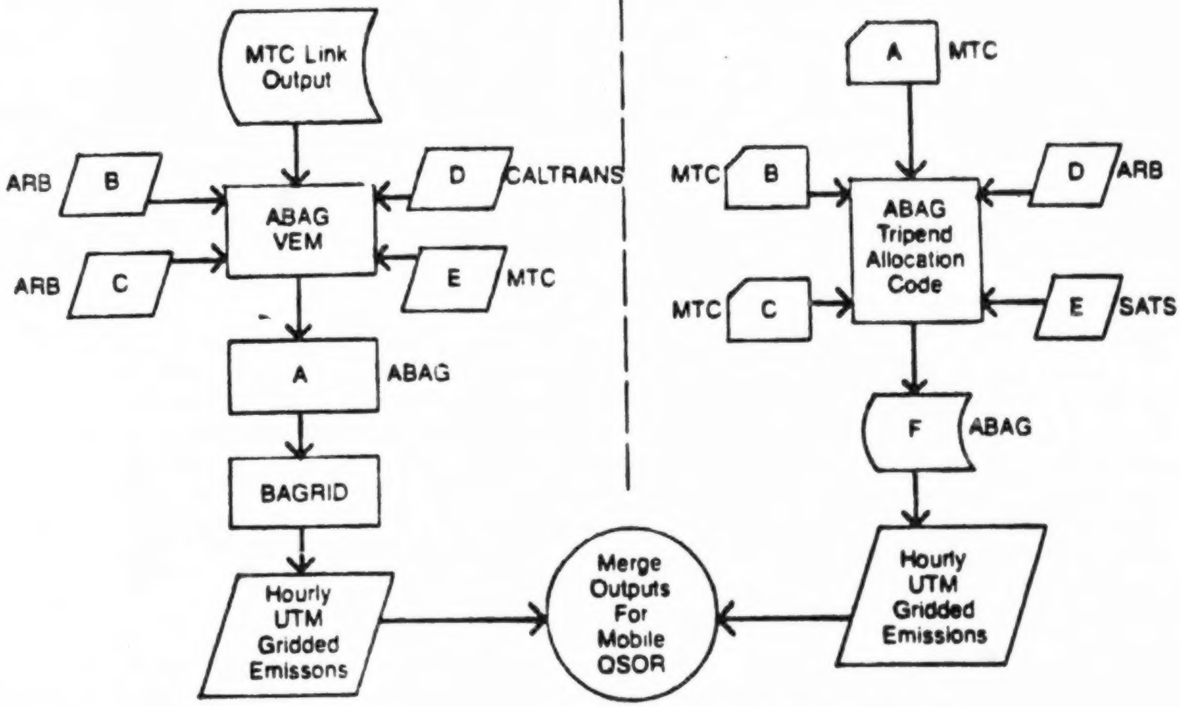
**These factors were subsequently adjusted to incorporate EPA's draft Supplement 8 factors (June 1977). The latest revision embodied in EPA's Mobile 1 (March 1978) was not available in time to be used in the analysis.

Figure C-2

Organization of the motor vehicle emissions code

A LINK EMISSIONS COMPONENT

B. TRIP-END EMISSIONS COMPONENT


MTC Link Output
ARB
B
D
CALTRANS
ABAG VEM
ARB
C
E
MTC
A
ABAG
BAGRID
Hourly UTM Gridded Emissons
Merge Outputs For Mobile QSOR
A
MTC
MTC
B
ABAG Tripend Allocation Code
D
ARB
MTC
C
E
SATS
F
ABAG
Hourly UTM Gridded Emissions

LINK EMISSIONS

Required input data

A — State plane/UTM transform

B — Emission and deterioration factors for each model year for 1975, 1985, 2000

C — Motorcycle emission factors, SO_2 and particulate emission factors for all vehicles, weighted average for 1975, 1985, 2000.

D — Updated speed correction equations for LDV, HDV, diesel.

E — Percentage truck and motorcycle VMT by hour and functional road type

TRIP-END EMISSIONS

Required input data:

A — Origin-destination trip tables for each travel model run (including intrazonal)

B — Hourly distribution of trip starts by trip purpose for four soak periods

C — Intrazonal VMT per zone

D — Cold start, hot soak emission factors for 1975, 1985, 2000 (weighted average over vehicle population)

E — Hot-soak period distribution

F — 440 zone/1 km grid conversion

ABAG = Association of Bay Area Governments
ABAGVEM = name of computer code with vehicle emission factors
ARB = California Air Resources Board
BAGRID = computer code to distribute link emissions to grid squares
CALTRANS = California Department of Transportation
LIRAQ = Livermore Air Quality Model
MTC = Metropolitan Transportation Commission
QSOR = name of source inventory file for LIRAQ model
SATS = Sacramento Area Transportation Study
UTM = Universal Transverse Mercator coordinate system

Finally, the modified SAPLSM code read the hourly and total daily link emissions for the entire network and performed the following:

- converted the plane coordinates of the MTC network to the UTM (Universal Transverse Mercator) coordinates required for LIRAQ.
- allocated the link hydrocarbon emissions into three LIRAQ-defined reactivity classes
- wrote an output file in a format suitable for input to LIRAQ

Trip End Emissions

A separate computer program for estimating trip end emissions produced hot start, cold start and hot soak emissions by zone and hour of day (see Figure C-3). It also computed hot stabilized emissions for intrazonal (i.e., within the zone) trips. The basis for the emissions computations was a trip table which was developed by the transportation model. From this table the program computed the number of trip starts (i.e., origins) and stops (i.e., destinations). The emissions rates were determined as a function of the parking (or shutdown) time before and after a trip (for the start and stop tripends, respectively) and whether the vehicle was catalyst or non-catalyst equipped.

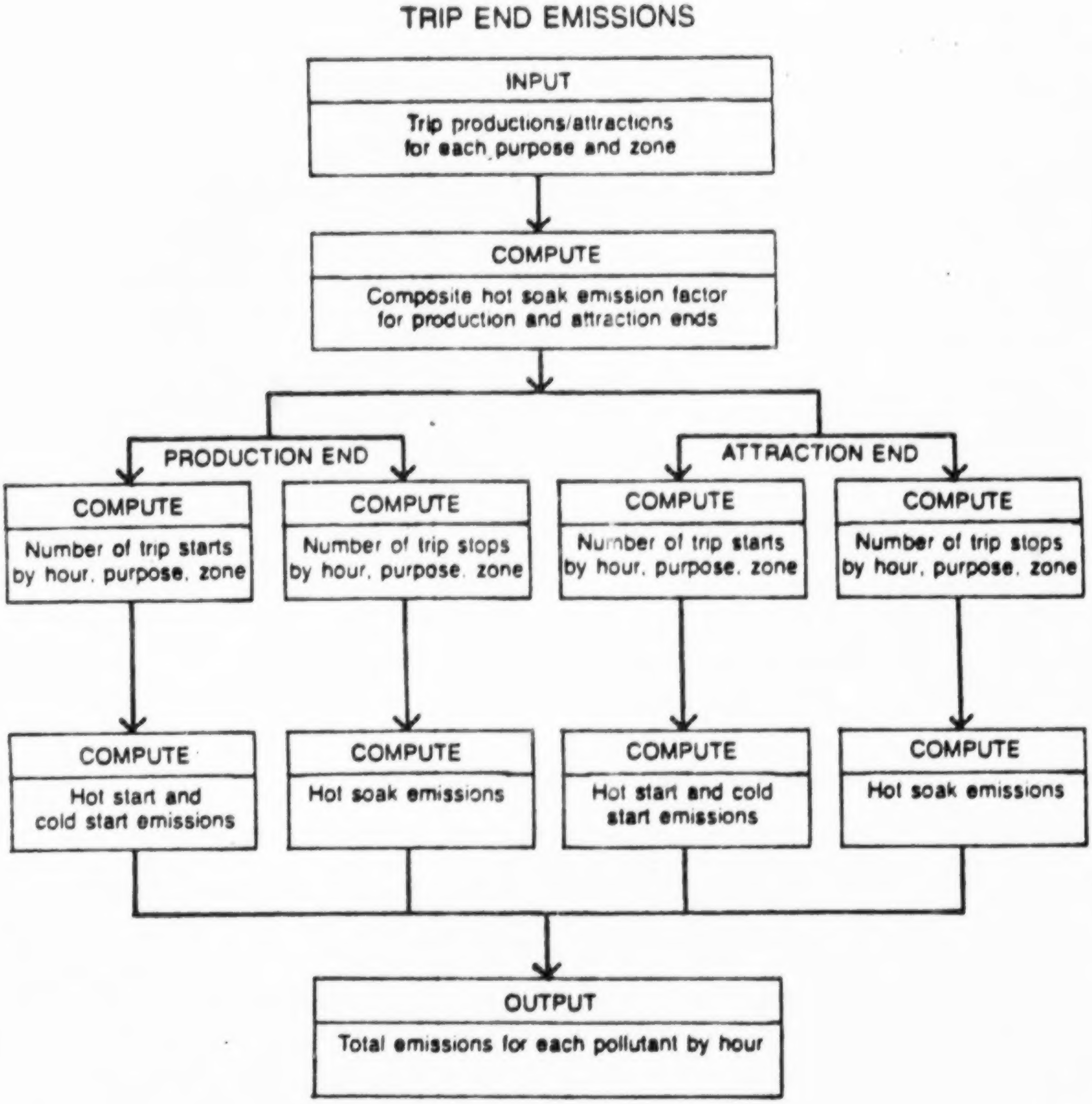
A special study by the California Department of Transportation provided parking time profiles by trip purpose and by trip origin or destination. From these profiles more accurate estimates were made of the percent of trip starts and stops experiencing hot and cold start and hot soak emissions. The trip purpose generally differentiated between the work-related and therefore long-term parkers and the non-work, and therefore, short-term parkers (e.g., shopping, recreational). The trip end identified whether the vehicle was starting or stopping at the home zone and thus the likelihood of a long parking period.

In computing trip end emissions, four trip types were considered. Considering work trips, for example, there are:

- the trip origins at the home (i.e., production) end, where many cold starts take place in the morning
- the trip destinations at the work (i.e., attraction) end, where many hot soaks take place in the morning
- the trip origins at the work (attraction) end, where many evening cold starts occur for the return trip
- the evening destinations back at the home (production) end, where hot soaks occur.

Note that the terms starts and origins, stops and destinations are used interchangeably.

Flowchart of trip-end emissions program.


TRIP END EMISSIONS
INPUT
Trip productions/attractions for each purpose and zone
COMPUTE
Composite hot soak emission factor for production and attraction ends
PRODUCTION END
ATTRACTION END
COMPUTE
Number of trip starts by hour, purpose, zone
COMPUTE
Number of trip stops by hour, purpose, zone
COMPUTE
Number of trip starts by hour, purpose, zone
COMPUTE
Number of trip stops by hour, purpose, zone
COMPUTE
Hot start and cold start emissions
COMPUTE
Hot soak emissions
COMPUTE
Hot start and cold start emissions
COMPUTE
Hot soak emissions
OUTPUT
Total emissions for each pollutant by hour

DENVER METROPOLITAN AREA

Stationary Sources

A general summary and comments on emissions inventory for the Denver study are listed in the following section. Table C-6 presents major assumptions made in estimating emissions from seven stationary sources categories.

- o Emissions data for stationary sources were compiled by the Colorado Department of Health.

- o Emissions data were collected by seven major source categories: point sources, gasoline service, solvent users, oil-based paint use, space heating, incinerators, airports.

- o Hourly emissions estimates were compiled on a 1 x 1 mile grid covering 900 square miles. However actual model runs were made with 2 x 2 mile grid.

- o Automobiles and point sources are major emission sources in the Denver Metropolitan region. Mobile sources account for about 80 percent of the total hydrocarbon emissions, 30 to 40 percent of the total NOx emissions, and 90 percent of the total CO emissions in the region. Point sources contribute approximately 15 to 20 percent of the hydrocarbon emissions and 40 to 50 percent of the NOx emissions.

- o SAI's Urban Airshed Model used in the Denver study requires emissions of reactive hydrocarbons and nitrogen oxides to determine the concentrations of photochemical oxidants. The reactive hydrocarbon emissions are further split into emissions of aldehydes, olefins, aromatics, and paraffins. The most recent Urban Airshed model chemistry is a 42-reaction "Carbon-bond" mechanism with four main classes of hydrocarbons, designated as single bonds (paraffins), fast double bonds (olefins except ethylene), slow double bonds (aromatics and ethylene), and carbonyl bonds (aldehydes and ketones).

- o The "Denver Model" treats two types of sources: "ground level" sources, whose emissions are injected into the surface layer of grid cells, and elevated point sources, whose emissions may be injected into any grid level depending on the calculated plume rise. Point sources having stack height less than 50 feet are combined with other ground level sources.

- o Hydrocarbon emissions are assumed to be 75 percent reactive and 25 percent nonreactive. Nitrogen oxide emissions are assumed to be 85 percent by weight nitric oxide and 15 percent nitrogen dioxide.

Method for Preparing the Motor Vehicle Emission Inventory in the Denver Metropolitan Area

In the Denver study, trip end emissions were lumped with hot-stabilized emissions. Mobile sources were separated into two files, "Auto Link" for traffic that crosses traffic zone boundaries and "Auto Area" for intrazonal traffic (traffic that remains within traffic zone boundaries). The link file was created by the Colorado Division of Highways using their transportation models. For the years 1974 and 1975 the inputs to the model were derived from traffic count data collected in 1971. For future years, traffic loadings were estimated from the Joint Regional Planning Program year 2000 land use plan, the Colorado Dept. of Highways road projections and the Denver Regional Council of Governments Empiric Activity Allocation Model for estimating types of trips. Vehicle emission factors were estimated using EPA AP-42, Supplement 5, predecessor of MOBILE1. A composite emission factor was used which combined trip end emissions with hot stabilized link emissions.

Since the emission of pollutants from automobiles and other traffic are dependent on the typical cycles of operation, which vary with location and type of roadway, the Division of Highways classified the links in the Denver transportation system under eight roadway types and four area types. For the years 1985 to 2000, however, only five different roadway types were used in this classification system. Depending on whether the operating conditions are peak (rush-hour) or off-peak, the Division of Highways estimated average vehicle operating speeds for each of the roadway types within each area type, as well as the diurnal variation in traffic flow for the Denver metropolitan region.

Using the link positions and lengths (inputs to the transportation models), the estimated average daily traffic (output of the transportation models), the estimated emission factors and the speed tables for the various link types, the Colorado Division of Highways estimated the Auto Link emissions as a function of the time of day and location within the Denver Highway Planning coordinate system. These emissions were then apportioned to the appropriate 2 mile square grids defined for the Denver area for application of the SAI photochemical model.

For "Auto Area" emissions, traffic loadings and diurnal cycle were estimated based on the demographic and geographic characteristics of a grid square. Emission factors were also derived from EPA-AP-42, Supplement 5. Once the Auto Area emissions were calculated for each square and temporally resolved, they were added to the Auto Link emission inventory.

REFERENCES

Bay Area Air Quality Management District, "Base Year 1975 Emissions Inventory, Summary Report," Bay Area Air Quality Management District, San Francisco, California, 1976.

Bay Area Air Quality Management District, "Base Year 1975 Emission Inventory - Source Categories Methodologies," Bay Area Air Quality Management District, San Francisco, California, 1976.

Hoffman, S. R., E. Y. Leong, and R. Y. Wada, "Air Quality Plan Development in the San Francisco Bay Region: Integrating Complex Regional Models into the Decision-Making Process," presented at the American Institute of Planners 61st Annual Conference, September 27 - October 1, 1978, New Orleans, Louisiana.

E. Y. Leong and R. Y. Wada, "Emission Inventory Projections: Hindsight, Insight and Foresight," proceedings of the Air Pollution Control Associstion Specialty Conference on Emission Factors and Inventories, Anaheim, California, November 13-16, 1978.

Perardi, T. E., M. Y. Kim, E. Y. Leong and R. Y. Wada, "Preparation and Use of Spatially and Temporally Resolved Emission Inventories in the San Francisco Bay Region," Journal of the Air Pollution Control Association, Vol. 29, No. 4, pp. 358-364, April 1979.

U. S. Environmental Protection Agency, "Air Quality in Denver Metropolitan Region, 1974 - 2000," Environmental Protection Agency, Region VIII, Denver, Colorado, 1977.

U. S. Department of Transportation, Federal Highway Administration, "Special Area Analysis - Part 4. Special Area Pollution," prepared by L. R. Seiders, Comsis Corporation for FHWA Urban Planning Division, Washington, D. C., August 1973.

U. S. Department of Transportation, Federal Highway Administration, "SAPOLLUT/SAPLSM User's Guide," FHWA Urban Planning Division, HHP-23, Washington, D. C., October, 1976.